ORIGINAL PAPER

Cash and the Hidden Economy: Experimental Evidence on Fighting Tax Evasion in Small Business Transactions

Ho Fai Chan[1] · Uwe Dulleck[1,2,3] · Jonas Fooken[4] · Naomi Moy[1] · Benno Torgler[1,5]

Abstract

Increasing the tax compliance of self-employed business owners—particularly of trade-specific service providers such as those involved in construction and repair work—remains an ongoing challenge for tax authorities. From a compliance point of view, cash transactions are particularly problematic when services are paid for on the spot, as these exchanges are difficult to audit. We present experimental evidence testing ten different policy strategies rooted in the enforcement, service, and trust/social paradigms, in a setting that allows payment either via a transaction that directly reports income for tax collection purposes or in cash, where taxes are only collected on reported income. Our sample includes both a typical subject pool of students, as used in most previous studies, and non-students who are active within service industries characterised by the opportunity to engage in cash transactions. While our comparative results show that, for both student and non-student participants, interventions that rely on greater enforcement by the tax authority have the greatest effect on compliance in our cash economy setting, treatments involving cooperative elements may be similarly effective in enhancing tax compliance. Given their effectiveness, cooperative approaches should therefore be considered for addition to the policy mix if implemented at relatively low costs, making both carrot and stick approaches promising to increase compliance in an environment where cash-for-service payments offer a common benefit for small businesses and their customers from implicit collusion that enables tax evasion.

Keywords Tax compliance · Hidden economy · Cash transaction · Laboratory experiment · Service industry

JEL Classifications O17 · C91 · H25 · H26

Introduction

Although maintaining a high level of tax compliance is central to a well-functioning society, one crucial challenge for research in this area is that the inherently hidden nature of tax evasion makes data collection difficult. Increasing self-employed business owners' tax compliance remains an ongoing challenge for tax authorities, particularly with respect to trade-specific self-employed service providers who encounter frequent opportunities to deliver services that can be paid in cash (e.g. construction workers, electricians, and plumbers). Cash transactions are particularly problematic from a compliance point of view when services are paid for on the spot and leave no paper trail. Such exchanges are difficult to audit, allowing service providers (e.g. tradespeople, massage therapists, cleaners, removalists, etc.) and sole

✉ Benno Torgler
benno.torgler@qut.edu.au

Ho Fai Chan
hofai.chan@qut.edu.au

Uwe Dulleck
uwe.dulleck@qut.edu.au

Jonas Fooken
j.fooken@uq.edu.au

Naomi Moy
n.moy@qut.edu.au

[1] Centre for Behavioural Economics, Society and Technology and School of Economics and Finance, Queensland University of Technology, Brisbane, Australia

[2] Crawford School of Public Policy, Australian National University, Canberra, Australia

[3] CESifo, Munich, Germany

[4] Centre for the Business and Economics of Health, University of Queensland, St Lucia, Australia

[5] CREMA – Centre for Research in Economics, Management, and the Arts, Zurich, Switzerland

traders to conceal income. We focus on building, renovating, and repairing services, a sector with particularly frequent opportunities for evasion and the ability to underreport cash transactions. Data from several OECD countries suggest that the building, renovating, and repairing sector is responsible for almost 50% of illicit work (Schneider & Enste, 2002).

Enforcement certainly matters in enhancing compliance, as it affects the financial considerations that motivate—at least in part—an individual's compliance choice. Yet, evidence has shown that it is not only the economic consequences of punishment that cause individuals to pay taxes (e.g. Alm et al., 1992; Torgler, 2002, 2007). Recently, the understanding of individual choice processes has been expanded by introducing aspects of behaviour or motivation that can be classified under the general rubric of "behavioural economics" or "behavioural taxation" (Torgler, 2022). This is broadly defined as an approach that uses methods and evidence from a variety of social sciences to inform the analysis of individual and group decision making. It is therefore valuable to explore potential instruments by applying a more complete understanding of individual (and group) decisions and one that is more consistent with empirical evidence. Behavioural economics has demonstrated that many individuals are motivated by social norms and intrinsic motivation, and that individuals are capable of learning social norms (Ostrom, 2005; Torgler, 2007). Erard and Feinstein (1994), for example, propose a model that describes the role of moral sentiments in tax compliance. Deviation from this norm creates psychological costs, which individuals try to avoid, as they are generally cooperative and are trying to make the pro-social and cooperative choice if the cost of doing so is not prohibitive. Dulleck et al. (2016) use a physiological marker to find support for the proposition that psychic stress is correlated with tax compliance. Research in tax compliance has convincingly argued that successful tax collection is not only the exercise of power (Alm & Torgler, 2011; Alm et al., 2010; Kirchler, 2007; Torgler, 2007), but that tax compliance, like much of human behaviour and institutions, is composed of a mixture of "love" and "fear" (Boulding, 1981).

More recently, however, both researchers and tax administrations have placed more emphasis on integrating the "love" or cooperative aspect, especially given that citizens' consent to pay taxes reflects identification with the taxing authority's objectives (Boulding, 1981). The view that both enforcement and cooperation are important is also reflected in central theoretical work that models tax compliance decisions (see, e.g. Kirchler et al., 2008; Alm & Torgler, 2011). Alm and Torgler (2011) suggest three core compliance paradigms: namely, the traditional "enforcement" paradigm; the "service" paradigm that recognises the role of the tax administration as a service provider to the taxpayers; and the "trust/social" paradigm that demonstrates the importance of ethics, trust, morality, and social norms in tax compliance (Alm & Torgler, 2011).

In our laboratory experiments, we test several interventions that reflect all three paradigms—*enforcement*, *service*, and *trust/social*—and have the potential to translate into real-life policies or strategies. While some interventions have received significant attention (e.g. increase in audit, moral suasion), others have not been explored in prior experiments and in sufficient detail (e.g. the provision of positive feedback, the possibility to show remorse after being audited). The comparative advantage of our study is to see the *relative* strengths of a large number of strategies evaluated concurrently, covering all three paradigms. Most experimental studies rely on one paradigm and only on a limited number of policies, which makes a comparative approach problematic, and makes it challenging to compare the insights from different studies that differ in several parameters and vary in their design, framing, or sample sizes.

Discussing the current state of knowledge about tax compliance behaviour based on laboratory experiments, Alm and Malézieux (2021) note that "despite the wide use of TEGs [Tax Evasion Games], it is disappointing—and surprising—that the impacts of many variables examined in these TEGs remain unclear" (p. 700). The common approach for laboratory experiments in compliance research is that, at the beginning of each round, each subject is given or earns income and then must decide how much income to report. Taxes are then paid at some rate on the reported income but not on the unreported amount, while the under-reporting is discovered and then penalised with some probability. The researchers then introduce policy changes such as changes in deterrence, public good provision, or other institutions (Alm 2019).

Our contribution is different to the one provided in these prior laboratory experiments, as we study the cash economy setting. Contrary to standard tax compliance experiments, most of the decisions in the cash economy are made in a social context where non-compliance decisions are a complicit act between multiple actors that enable evasion. Specifically, we examine the role that a second actor—the buyer—has in a seller's willingness to be non-compliant regarding the business' income, thereby providing insights into how different policies work in a cash economy environment. In our experiments, the additional actor, the consumer, allows the possibility of a cash transaction, opening the possibility for tax evasion. The presence of a potentially complicit buyer can influence the seller's ultimate compliance in various ways. It may influence the seller's perception that the buyer will be willing to enable evasion if asked to do so, it may affect the buyer's actual willingness to enable evasion and can influence the seller's rate and amount of compliance

once given the opportunity to evade. All these factors combine to a final decision to pay taxes on income received.

By examining tax evasion within business interactions rather than personal income tax evasion—as is most commonly studied—we can incorporate the effect that a second actor has in enabling tax evasion. There is currently only a very limited empirically informed understanding of what happens if another actor is involved in the evasion decision. It may be that moral costs of non-compliance decrease when adding another actor (e.g. via sharing the blame). Or it could be that introducing a stronger social context supports compliance via stronger influences of social norms. In addition, our results contribute to the literature on the cash and shadow economy, as previous research has mainly relied on macro data rather than experimental micro data to derive empirical insights and derive policy implications.[1]

Furthermore, we contribute to research on the use of non-student participants in tax compliance experiments, which have been analysed as subject pools in field experiments (e.g. Bott et al., 2020; Hallsworth, 2014; Torgler, 2004, 2013), surveys (Górecki & Letki, 2021), and online experiments (Farrar et al., 2019) but have not been widely studied in the laboratory (for exceptions, see, e.g. Gërxhani & Schram, 2006; Alm et al., 2012; Choo et al., 2016). Alm and Malézieux's (2021) meta-study stresses that laboratory experiments "are most often run with student subjects, which is often cited as the main concern about the external validity of those experiments" (p. 706). Given the experimental design exploring the relationship between a house owner and a tradesperson—and given the large construction industry in Australia (the location of our study) is comprised mainly of sole traders and provides many opportunities for cash transactions—we recruited non-students working in trades as participants. Using tradespeople as subjects is particularly interesting as they are likely to face decisions about whether to act in the cash economy when conducting their daily jobs.

However, a shortcoming of using a non-student sample of tradespeople alone is that it makes our study less comparable to previous studies. For comparison purposes, we therefore conducted two laboratory experiments, one with students and one with non-students. Considering the prior literature, we take a two-pronged approach by testing the hypotheses with students and then testing the hypothesis with non-students. In doing so, we avoid the potential noise that is created by switching between hypotheses and subject pools or when comparing individual studies that rely on only one subject pool. This allows us to test a set of instruments that may be available to tax authorities for understanding tax compliance behaviour in a situation where large cash transactions are possible. Thus, our results provide new insight into what types of interventions can influence the intent to be compliant with tax directives within the cash-based economy.

In our results, in line with existing literature on tax compliance with respect to personal income taxes, we observe that traditional *enforcement* approaches, such as raising audit rates or endogenous audits are highly effective in reducing tax evading behaviour in cash transactions. However, we also find that interventions appealing to the *service* and *trust/social* compliance paradigms are effective in increasing tax compliance, especially with the standard (student) participant pool. Given that enforcement is known to increase compliance, we also test our *service* and *trust/social* treatments relative to the effect of the *enforcement* treatments, showing that (with one exception) the non-*enforcement* treatments do not result in a significantly smaller effect than the *enforcement* treatments. In particular, the provision of assistance in tax declaration and activation of norm-based peer effects are useful strategies for increasing compliance and reducing loss in tax revenue due to the cash economy. Options to pre-fill tax declarations also provide large positive effects on compliance, despite the financial cost associated with such a service. Overall, our comparative analysis on the effectiveness of a rich set of tax policy interventions demonstrates the significant potential for non-audit-based measures in tackling tax evasion in the cash economy. Given the large financial cost associated with increasing audit rates, we provide useful empirical evidence for tax departments who wish to consider using a more cost-effective method to increase compliance within the hidden economy.

Theoretical Framework, Experimental Treatments, and Hypotheses

Figure 1 summarises our theoretical framework. We consider a series of options and instruments that a tax administration could use to combat tax evasion and promote tax compliance. These strategies can be classified into three core paradigms: the enforcement paradigm, the service paradigm, and the trust/social paradigm (Alm & Torgler, 2011). The arrows in Fig. 1 indicate that a respectful and legitimate authority engaged with its citizens will move away from placing more coercive enforcement at the top of the pyramid to achieve compliance (Braithwaite, 2007) and shift its approach towards alternative paradigms such as the service or trust/social paradigms. Thus, the two non-enforcement paradigms offer an additional set of potential instruments. Such developments have been observed by tax administrations worldwide who have been moving from a largely "stick"-based approach with a heavy focus on enforcement, towards a more cooperative model

---

[1] For exceptions see, e.g. Maciejovsky et al. (2012) or Chan and Song (2021).

Fig. 1 Theoretical framework
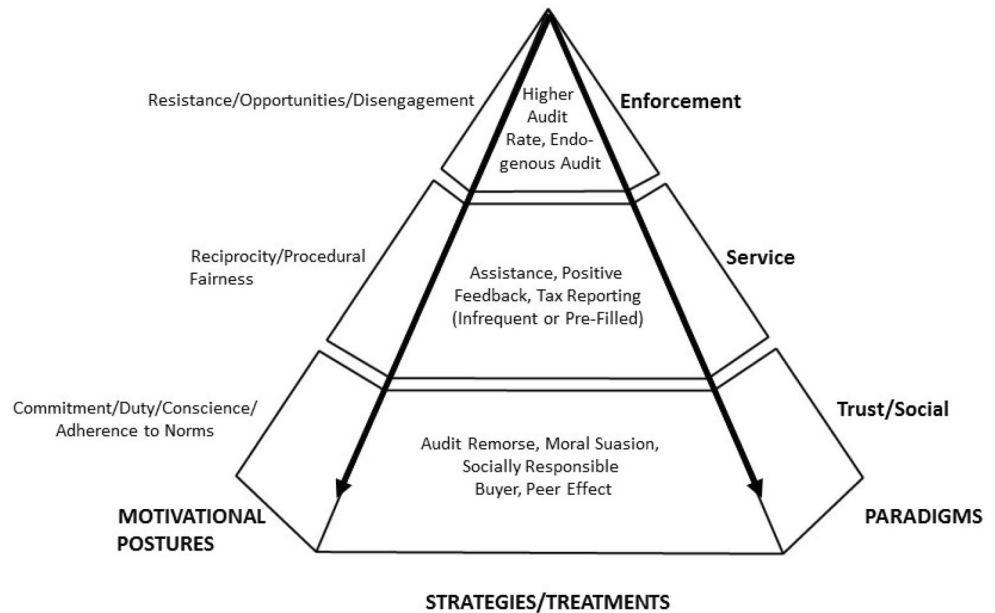
("carrot" approach) that considers service and trust elements. As Braithwaite (2007) stresses, "de-escalation is desirable, once cooperation is forthcoming" (p. 5). Tax authorities that are seen as part of the same community as the taxpayers (Kirchler et al., 2008) can promote procedural fairness, which enhances the feeling of reciprocity and affects social exchange in a positive way. Putting more emphasis on instruments that show trust in the taxpayers promotes commitment, obligation, and social responsibility. The main advantage of enhancing voluntary cooperation is that it activates a stronger sense of duty. As Kirchler et al. (2008) point out, "an increase in trust can increase the power of authorities because citizens support the tax officers and ease their work" (p. 213). Slemrod (2019) emphasises that the "leading alternative, but not mutually exclusive, paradigm to the Becker–Allingham–Sandmo model is one that relies on duty, conscience, and adherence to norms" (p. 947), which are core motivational postures in the trust/social norm paradigm.

In our experiment, we use 10 treatments that incorporate important examples of all three paradigms. Our treatments that build on the enforcement paradigm use (1) higher audit rates and (2) 'endogenous' audit rates that increase when an individual has been found non-compliant in the past. Our treatments that build on the service paradigm use (3) the offer of assistance to fill in the tax declaration, (4) positive feedback on decisions to pay taxes, and reporting aspects such as (5) the use of greater taxpayer autonomy by allowing for less frequent reporting, and (6) less frequent reporting with the option to have the tax declaration form pre-populated. Our treatments that build on the trust/social paradigm use (7) the option to correct the declaration after non-compliance is detected but with a greater penalty when caught

being non-compliant again (audit remorse), (8) moral suasion referring to the social norm of paying taxes, (9) appeals to the social responsibility of (non-tax-paying) buyers in avoiding cash transactions (socially responsible buyer), and (10) information on the average level of tax declared by other taxpayers (peer effect). We note that the service and the trust/social paradigm are intricately connected. For example, providing taxpayers with more autonomy by allowing less frequent reporting can be seen as a sign of trust. We employ all treatments in our experiments with students and treatments 1, 3, 5, 8, and 10 in our experiments with non-students (i.e. tradespeople). We discuss the three paradigms, how our treatments are embedded in them, and what hypotheses are made based on the theory that motivates them in further detail below.

Enforcement Paradigm

The enforcement (or deterrence) paradigm is based on the standard economic model, inspired by Gary Becker's (1968) economics of crime approach. The model stipulates that selfish–rational decision makers will try to pay as little tax as possible, as the marginal utility from tax payments (e.g. through the receipt of publicly provided services) is lower than the marginal utility of directly using foregone tax payments for one's own use (e.g. through consumption). The model by Allingham and Sandmo (1972) often serves as the theoretical benchmark in this approach and predicts that selfish–rational taxpayers will not pay any taxes unless there is a risk of non-compliance being detected and punished; and even in this case they will continue to evade, depending on their willingness to take risks, the probability of detection, and the size of the fine for evasion. In this model, the only

way to increase compliance is to increase the (perceived) risk that non-compliance is detected and to impose higher fines in case of detection. In such a traditional enforcement paradigm, taxpayers are therefore seen and treated as potential criminals (Alm & Torgler, 2011). As Braithwaite (2007) points out, the command-and-control operational system tries "to accomplish their mission of catching "the scoundrels" who do not pay their tax" (p. 4). We chose two of the core enforcement-based approaches and instruments to reduce tax evasion: the use of (1) higher audit rates and (2) 'endogenous' audit rates that increase when an individual has been found non-compliant in the past.

Higher Audit Rates

Traditionally, audit rates have been a key variable of analysis, particularly among economists, because they reflect the fact that non-compliance is a risky decision that may lead to detection and (costly) punishment. That is, the higher the probability of audit, the less likely a taxpayer will evade taxes (Alm & Malézieux, 2021). Evidence from lab experiments indicates that a higher audit rate leads to more compliance (Alm, 1999; Blackwell, 2010; Torgler, 2002), with one study indicating that audit rates are successful in improving compliance in cash transactions (Chan & Song, 2021). Additional evidence has suggested that higher auditing rates induce probability neglect, where the individual receives the information, overweighs the probability of an audit occurring, and is subsequently more compliant (Bérgolo et al., 2017). We therefore hypothesise that higher audit rates increase compliance in a business setting involving cash transactions to check whether our data in our setting confirms previous research results.

Hypothesis$_{(Higher\ Audit\ Rates)}$ Higher audit rates *increase* compliance relative to the baseline in cash transactions.

Endogenous Audit

Many tax compliance experiments have integrated endogenous audit selection rules to increase external validity, as tax agencies often do not select tax returns randomly for audit but instead use information from the returns to determine audits (Torgler, 2002). For example, the Internal Revenue Service (IRS) uses the Discriminant Index Function (DIF) formula based on items reported on current tax returns in its selection of audits (Alm et al., 1993). Other countries follow a similar practice (Roth et al., 1989). Thus, the probability of audit is endogenous in the sense that it depends on the behaviour of taxpayers. There are different ways of introducing selection rules. A common one is to use past observed behaviour, increasing the audit probabilities for those taxpayers known to have been non-compliant in the

past, as used in our treatment, while others target audits on taxpayers who report less than some cut-off level of income. Experimental results indicate that endogenous audit rules can generate significantly greater compliance than random audit rules (Torgler, 2002) and we expect the same in our data.

Hypothesis $_{(Endogenous\ Audit)}$ Endogenous audit rules *increase* compliance relative to the baseline in cash transactions.

Service Paradigm

A service paradigm emphasises the role of the tax administration in facilitating and providing services to citizens; as such, policies are introduced that assist taxpayers in filling tax returns and paying taxes (Alm & Torgler, 2011). This developed out of the New Public Management of the 1980s that required public administrations to be more service-oriented (Torgler & Murphy, 2004). Taxpayers are no longer seen as criminals but rather as clients (Alm & Torgler, 2011). Thus, the tax administration aims to deliver higher quality services to taxpayers. The tax administration is therefore responsive to the conduct of the taxpayers. Evidence indicates that friendly and respectful treatments and a greater service orientation enhance tax compliance and tax morale (Feld and Frey 2002, Torgler & Murphy, 2004, Torgler et al. 2008).

Assistance

Assistance to taxpayers encourages identification with the task (i.e. compliance) and with the entity offering assistance (i.e. the tax collector in this case) which can promote reciprocity through increased mutual obligations. Humans tend to repay what others have provided to them, as they feel a duty to reciprocate favours (Boulding, 1981), a deeply rooted adaptive mechanism in human nature (Cialdini, 2007). Such reciprocity can contribute to the development of integrative structures between taxpayers and the tax administration and failing to comply may trigger internal discomfort and psychological costs (Erard & Feinstein, 1994). Mazar et al., (2008, p. 634) stress that if "a person fails to comply with his or her internal standards for honesty, he or she will need to negatively update his or her self-concept, which is aversive. Conversely, if a person complies with his or her internal standards, he or she avoids such negative updating and maintains his or her positive self-view in terms of being an honest person". Mazar et al. (2008) also argue that people will comply with their internal standards even if compliance necessitates an investment of effort and sacrifice.

The offer to help with the tax compliance decision primarily aims to signal to taxpayers that the tax authority is

cooperative and therefore enhances the attitude towards the tax authority. Torgler et al. (2008) use data from the US and Turkey to analyse interactions between taxpayers and the tax administration, finding that positive attitudes towards the tax authority (e.g. how taxpayers rated tax administrations' job, their honesty and fairness, and their helping and information behaviour) significantly increase tax morale. A respectful and fair treatment of taxpayers also induces respect for the tax system and therefore promotes cooperation (Smith, 1992). These findings are echoed by recent studies where interactional fairness positively influences tax compliance, and the provision of information services by the tax authority decreases tax evasion (Farrar et al., 2019; McKee et al., 2018). We therefore hypothesise that the provision of assistance in submitting their tax declaration increases compliance in cash transactions, as taxpayers reciprocate the cooperative help they receive.

Hypothesis (Assistance) Assistance *increases* compliance relative to the baseline in cash transactions.

Positive Feedback

Offering positive feedback for being compliant—via something as simple as a thank-you note—is intended to motivate and reward desired behaviour. Instead of raising the relative psychological cost of not paying taxes, the instrument of reward raises the psychological benefits of paying taxes (Feld et al., 2006). Rewards are widely used in daily business activities and in society in general, as an acknowledgement of the desired compliance or behaviour. A thank-you note might be perceived as supportive, bolstering future compliance, and strengthening the attractiveness of rewarding "good" taxpayers, again primarily through signalling reciprocity from the side of the tax collector. It also reinforces the social norm of compliance and communicates the cooperative nature of the tax authority. Currently, we have limited empirical information on whether a thank-you note for full compliance supports compliance. Taxpayers may be more willing to comply and be cooperative towards another reciprocal decision maker. Nevertheless, whether such treatment indeed increases compliance is an empirical question. The power of rewards in shaping human behaviour has long been a topic among social psychologists (see, e.g. Thorndike, 1911, 1932; Skinner 1938, 1953; Nuttin & Greenwald, 1968). Rewards are expected to change the relative prices such that paying taxes becomes a more attractive alternative to evading taxes, but this does not necessarily mean that the effect is large enough to enhance compliance. It may produce sustainable compliance among generally honest taxpayers, but taxpayers who are able to be non-compliant may not reciprocate to a positive acknowledgement such as a thank-you note.

Hypothesis (Positive Feedback) Positive feedback *increases* compliance for cash transactions relative to the baseline.

Infrequent Reporting

The requirements regarding the frequency of income reporting and paying taxes can differ between different types of taxpayers (e.g. personal income tax is reported and paid yearly, but a share of income is withheld for every pay period, whereas freelancers/sole traders pay estimated taxes in quarterly "Pay As You Go (PAYG)" instalments that they can request to be adjusted if the estimate is inaccurate for the quarter). The reporting and payment frequency between different forms of income often creates complex taxation environments resulting in directives that are used to simplify the tax system, for example, a realignment of reporting periods. However, the type of realignment (e.g. a reduction or an increase of the reporting frequency) could have a varying impact on tax compliance. Reducing the number of reporting periods, in practice, may allow taxpayers to smooth their underreported income, whereas more frequent tax reporting should make it more laborious to misreport income when continuing to operate consistent bookkeeping. On the other hand, the lower frequency of tax reporting may reduce a seller's perception of the probability of being detected and thus provide a greater opportunity for evasion. Furthermore, when reporting income for multiple periods, taxpayers may (deliberately or unintentionally) 'forget' income to report, reducing psychological costs of evasion. Together these elements suggest that tax compliance should be lower when income is reported less frequently. However, infrequent reporting may have a compliance-increasing effect if it allows the tax authority to signal cooperativeness with the taxpayer and trust in the taxpayer's honesty. As real-life bookkeeping factors are largely abstracted away in the experiment and reporting periods were close together, allowing little moral wiggle room to forget income, we expect a small positive effect on compliance under infrequent reporting.

Hypothesis (Infrequent Reporting) Infrequent reporting *increases* compliance relative to the baseline in cash transactions.

Infrequent Pre-Filled Reporting

Pre-filled reports provide a method of assisting taxpayers, reducing the transaction costs and uncertainty or even anxiety costs of the taxpayer. The use of pre-populated tax returns is expected to increase the compliance of taxpayers, through the removal of self-reporting-based errors. Several studies show that the average level of compliance is higher for pre-filled returns (Doxey et al., 2021; Fochmann et al., 2021; van Dijk et al., 2020). The caveat, however, is that

compliance for pre-filled is only higher when the undocumented income is included, or when the estimated income level is correct or greater than the actual income (Doxey et al., 2021; Fochmann et al., 2021; van Dijk et al., 2020). In the experimental context, pre-filled tax returns may appear as a welcome service that reduces mental effort and possibly even alleviates uncertainty around the decision to report the correct or incorrect amount for tax payments, both of which may induce psychological costs. We, therefore, hypothesise that offering the option to correctly pre-fill the tax return leads to higher compliance in cash transactions.

Hypothesis (Pre-Filled Infrequent Reporting) Pre-filled infrequent reporting *increases* compliance in cash transactions relative to the baseline and relative to the infrequent treatment.

Trust/Social Paradigm

Like the service paradigm, the trust/social paradigm is based on the idea that a trustful relationship between the tax authority and the taxpayer will support compliance. Trusted taxpayers may reciprocate the trust in them by being more committed and therefore compliant, facilitating the work of the tax administration (Kirchler et al., 2008).

Audit Remorse

Where a tax authority offers taxpayers the opportunity to reconsider their declaration in cases of misreporting, it often relies on their remorse but also increases the punishment when the adjusted reporting is still incorrect. There are two mechanisms at work here: the first is that it signals to the taxpayer that the tax authority treats them with respect and fairness—acknowledging that mistakes can happen—which may generate respect for the tax system and may lead to a higher level of cooperation. Some people become non-compliant by mistake and such individuals are usually willing to correct their behaviour and transform into compliant taxpayers. Thus, informing taxpayers about their non-compliance and allowing them to correct their behaviour offers the chance to integrate accidental non-compliers into the taxation system. It may provide a signal to taxpayers that they are expected to comply in the future, as it includes a higher penalty when non-compliance is detected a second time, which builds on the standard economics deterrence model. Furthermore, offering the option to reconsider the tax declaration provides the taxpayer with more procedural information (audit feedback without punishment), increasing the potential capacity for cooperation and commitment (Ostrom, 2005).

We are not aware of a previous experimental study that has investigated audit remorse. To predict the effect of the audit remorse treatments, it is useful to not only think about procedural fairness and reciprocity, but also about how compliance processes are linked to trust, motivation, and commitment (Torgler & Schneider, 2009), and balancing the concepts of trust and power (Batrancea et al., 2019; Kirchler et al., 2008). Such a structure of audit remorse signals that if the proffered trust is not reciprocated, harsher consequences are used. There is evidence that intensification of enforcement efforts is a successful strategy for increasing tax compliance after a tax amnesty (Alm et al., 1990). It might be seen as a fair warning, especially for those taxpayers who were honest before the tax amnesty; the goal is to convince tax delinquents that tax evasion is morally wrong (Fisher et al., 1989). We therefore hypothesise that for cash transactions, compliance would be greater in the audit remorse treatment.

Hypothesis (Audit Remorse) Audit remorse *increases* compliance relative to the baseline in cash transactions.

Moral Suasion

The mechanism through which moral suasion operates is by describing a prescriptive norm (e.g. what is right or wrong) to decision makers and relying on individuals to follow this norm—which can be viewed as indirect reciprocity. Economists may be sceptical about the effects of moral suasion, particularly in the long term or in competitive environments (Torgler, 2004), but social psychologists have demonstrated the power of moral suasion or moral appeals (see, e.g. Cialdini's, 2007 seminal work on persuasion), such as increasing the salience of how tax money is spent (e.g. charity). Research in marketing relies heavily on persuasion as a tool to influence human behaviour, as the goal of marketing is to form and change attitudes and actions (Torgler, 2013). Less evidence is available on how moral suasion or moral appeals shape tax compliance (Torgler, 2004, 2013). Field experimental evidence provides some (still limited) support for the proposition that moral suasion matters, reporting barely any effect on tax compliance when used at the local level, where moral suasion might be most effective (Blumenthal et al., 2001; Torgler, 2004, 2013). However, there is no evidence on whether it would affect compliance when a cash transaction and a second actor (the buyer) is involved. We hypothesise that tax compliance is higher under moral suasion than in the baseline when cash transactions occur.

Hypothesis (Moral Suasion) Moral suasion *increases* tax compliance in cash transactions relative to the baseline.

Socially Responsible Buyer

While tax compliance is determined by the seller, buyers play a role by providing the opportunity of evasion when

they implicitly collude with sellers by accepting cash transactions. Yet, the responsibility for making an honest decision about tax compliance is relegated to the seller. In fact, research indicates that participation in the cash economy is often initiated by buyers in transactions (Horodnic et al., 2021; Williams & Kosta, 2020). Reminding buyers of their role highlights the social norm of compliance and the effect that buyers can have by rejecting cash offers, which has an indirect effect on compliance levels. In other words, if buyers refuse cash offers, the sellers are limited in their ability to evade taxes. Thus, we hypothesise that this indirect effect can increase tax compliance levels.

Hypothesis (Socially Responsible Buyer) Reminding buyers of their role *increases* tax compliance levels relative to the baseline within the cash economy.

Peer Effects Seller

There is substantial evidence that peer effects matter for tax compliance (Frey & Torgler, 2007; Spicer & Becker, 1980; Webley et al., 1985), similar to other illegal or non-compliant activities such as assassinations, hijackings, corruption, kidnappings, serial murders, and littering (Bikhchandani et al., 1998; Dong et al., 2012; Torgler et al., 2009). Perceived social norms of compliant behaviour drive such peer effects, as does the risk of getting caught. Kahan (1998) suggests that the decision to commit crimes in general is highly interdependent, based on the perceived behaviour of others: "When they perceive that many of their peers are committing crimes, individuals infer that the odds of escaping punishment are high, and the stigma of criminality is low. To the extent that many persons simultaneously draw these inferences and act on them, moreover, their perceptions become a self-fulfilling reality" (p. 394). Hence, there is both a norm-based effect and one through changing the perceived probability of detection. We therefore hypothesise that compliance would be higher when higher peer compliance was communicated to participants.

Hypothesis (Peer Effects Seller) Communicating to participants that their peers are generally tax compliant *increases* compliance in cash transactions.

The Strength of Different Paradigms

The effectiveness of the enforcement paradigm is well established, however, the service-based and trust/social-based approaches have continued to be investigated or suggested as potential alternatives even though their effectiveness is less universally observed or identified in prior research. The reason for this continued pursuit is that these more behavioural economic approaches to compliance typically incur

significantly lower costs, making them attractive to policy makers and the tax administration. We therefore test the relative effectiveness of enforcement-based approaches to the service-based and trust/social-based approaches, grouping treatments 1 and 2, 3 to 5, and 6 to 10 into three broader approaches. Having previously hypothesised a compliance-increasing effect for all our approaches, we predict that all paradigms increase compliance and that the non-enforcement approaches have a similar effect on compliance as the enforcement approaches do.

Hypothesis (Paradigms) The enforcement and the non-enforcement approaches do not significantly differ in their effectiveness to increase compliance.

Methods

Experimental Design and Procedures

Due to the difficulty in collecting primary data on tax evasion, laboratory experiments are an essential tool in tax compliance as researchers strive to generate their own data (for an overview, see Andreoni et al., 1998; Alm, 1999, 2012; Torgler, 2002). The beauty of this approach lies in the ability to experimentally test researcher interests while isolating the effect under exploration. Experiments can unveil otherwise latent phenomena by collecting data on counterfactual cases that may not be observable in reality. Furthermore, other scientists can replicate the experimental conditions. Thus, it is not surprising that we have observed an increasing number of laboratory experiments since the 1990s (Torgler, 2002, 2016). To address the question of tax compliance by small business owners in the presence of different tax enforcement regimes, we use two experiments. As the basic framework of interaction is the same for both experiments, the common features are described first before describing how the two experiments differ from each other.

In both experiments, we used a framework that describes a stylised interaction between service providers (sellers) and customers (buyers).[2] Decisions were framed as a tax compliance decision in the context of sourcing services from a provider that may be paid in cash. Participants were told that they would make decisions in a service provider–customer framework and would interact with other participants in these roles throughout the experiment.

Upon arrival at the laboratory, participants were welcomed, and it was explained that they would make decisions throughout the experiment, for which they would be paid. This ensures that the decisions of participants are

---

[2] We use these terms interchangeably.

incentive-compatible—if they prefer to choose the dishonest declaration, they can do so and reap the economic benefits from this course of action. If they make an honest declaration instead, this comes at a real cost, as they receive lower monetary income at the end of the experiment.

After participants were familiarised with the rules that determine their payments, they received an explanation of how they would interact with others in the experiment. They were told that there would be two roles—service providers and customers—who would interact over several rounds. In each round, the customer needs to *get a job done*, hence, to receive a service to avoid losing 80 experimental dollars (from the 100 experimental dollars the customer receives each period). On the other side of the interaction, the service provider makes an offer to the customer for their service of *completing the job*. Participants were told that they could think of it as the relationship between a house owner and a tradesperson:

> To understand the interaction between service providers and consumers consider the consumer to be a house owner who needs some job to be done in the house which he cannot do himself. In real life this could be the repair of your swimming pool, of your hot water system or the refurbishment of a fence around your house. While these jobs may not need to be fixed straight away, there is a greater cost of ignoring the problem. The service provider in turn is someone who can do the job, such as a pool repair person, a plumber or someone specialised in fixing fences.

Furthermore, participants were told that the price paid by customers to the service provider was to be understood as income, that the service providers had to pay taxes on any income earned, and that the tax rate was 40%.

When service providers enter their offer to be submitted to the customer, they were asked if they wanted to include a cash option to the customer, and the customer was asked if they would accept a cash offer. Cash offers implied that the price to be paid by the customer would be 10% lower. Service providers and buyers were asked independently whether they wanted to accept the cash offer, and the cash offer was only implemented if both opted for the cash option. When the cash option was implemented, the income of the service provider (hence, the price offer minus 10%) was not automatically taxed. Instead, the service provider had to declare how much income had been received. Hence, in the case of cash transactions, it is possible for service providers to evade taxes by under-declaring income. By contrast, non-cash transactions are automatically 'declared' correctly (except in the infrequent treatments). Participants were informed that money collected as taxes would be paid to a university charity, which can be understood as a contribution to a public good, similar to tax revenue that is used for publicly

provided services. Participants were also told that after each declaration, there was a possibility that the service provider would be audited. The audit probability (in the baseline treatment) was set to 10%. If a service provider were audited and the amount declared was lower than the correct amount, the service provider would have to pay a fine, which corresponded to two times the underpaid amount.

After describing this general outline to participants of how they would interact with others in the experiment, participants were advised that there would be two parts of the experiment. Each part had six rounds. In every round, a different customer would interact with a different service provider. Participants were advised to read any instructions on later screens carefully, as these could include further information. In experimental terms, these further instructions represent the different treatments.

Before starting with the decision making, participants had to answer two control questions to ensure that they understood the game and how they would be paid based on their decisions. Control questions had to be answered correctly before participants were able to begin making decisions in the experiment. The full instructions for buyers and sellers—including instructions about potential audits—are reported in the Appendix (see Figures A1 to A5, A9 to A12).

The experiments include a *baseline* and ten treatment conditions that are classified into the three paradigms. In the *baseline* condition, participants received information about the experiment as outlined above with an audit probability of 10%. The following lists all treatments used in our experiment (see Appendix Figures A6 to A8).

Enforcement Paradigm

1. The *higher audit rate* treatment, which notified participants that in the following three periods the industry of service providers had come under special scrutiny, implying a doubling of the audit probability (relevant for the sellers).
2. The *endogenous audit* treatment, in which participants were informed that their personal future audit probability would be doubled if they were found to be non-compliant.

Service Paradigm

3. The *assistance* treatment, in which participants were informed that they could request further help on how to comply by asking a research assistant who was available for compliance questions. (Only occurs in Part 2 in the experiment).
4. The *positive feedback* treatment, which included a message of thanks to those participants who had made a fully compliant declaration. This message was not pro-

vided to non-compliant participants (independent of them being audited or not).

5. The *infrequent reporting* treatment, in which service providers had to report their income only after three periods as one large instalment (income history of the last three rounds was provided, see Figure A8).

6. The *infrequent pre-filled reporting* treatment, which was based on the infrequent condition. In this treatment, participants were given the opportunity to have their tax declaration pre-filled based on their income at a small cost (5 experimental dollars).

Trust/Social Paradigm

7. The *audit remorse* treatment, in which sellers who had been audited and identified as under-reporting were informed that the amount declared appeared too low. They were told that they had the opportunity to reconsider their declaration and were informed that the penalty would be tripled if they were caught under-reporting (again) after reconsidering.

8. The *moral suasion* treatment, in which both customers and service providers were reminded that tax money served a common good and that it would be paid to charity, namely a food bank at the hosting university. Therefore, it was pointed out that paying taxes in this experiment was important from a common good perspective. (Only occurs in Part 2 in the experiment).

9. The *socially responsible buyer* treatment, in which buyers were informed that accepting cash offers would provide sellers with the opportunity to evade taxes—and that buyers could consequently play a part in increasing compliance by refusing cash offers. (Only occurs in Part 2 in the experiment).

10. The *peer effects seller* treatment, in which sellers were informed that their declaration was *below/about/above* the industry average of declared income based on previous experiments. While not asking sellers to reconsider their declaration, they had the option to then adapt the amount declared.

Once the two parts of the experiment were completed (i.e. after the two rounds of 6 periods each), participants filled out a post-experimental questionnaire, which provided further demographic information and self-reported attitudes of participants. The full list of questions is included in the Appendix (Figure A13). Questions relating to tax compliance, demographic questions (e.g. gender, age, and nationality), and business ownership were compulsory, while others (e.g. income and religion) were voluntary.

As mentioned above, we used two experiments, one with students and one with tradespeople, to study tax compliance when there are opportunities for cash transactions.

Both experiments used the same experimental framework as previously mentioned but differed in some design features. Experiment 1, the study with students, used a *between-subjects* design to test the effect of different treatments, while Experiment 2, the study with tradespeople, used a *within-and-between-subjects* design. In addition, the amounts of money that could be earned in the two studies was higher for tradespeople. The details for Experiment 1 and Experiment 2 are described below.

Experiment 1

Experiment 1 was conducted in a computer laboratory at Queensland University of Technology in Brisbane, Australia, between 7 June and 24 June 2016. 266 volunteer student participants were recruited and played the role of a seller and of a buyer. 44% were female, the average age was 23.4 years, and 44.4% of participants were Australian nationals. The sample of participants consisted of undergraduate and postgraduate students (70.2% Bachelors, 20.9% Masters, and 8.9% PhD). We control for these characteristics as well as the nationality of participants because tax compliance has been found to differ between countries.

After going through the general instructions described in the experimental design and procedures, participants were randomly assigned the role of customer or service provider. Participants were informed that they would be either a service provider or a customer in the first part (of six rounds) and would switch to the other role in the second part (of six rounds). All participants were subjected to the *same* treatment for the set of rounds, in their differing roles. The treatments were introduced at the beginning of the rounds and were applied to all six rounds. Hence, participants were only presented with *one* treatment for each role, and because we only study the decision of service providers, in our analysis, all participants are only allocated to one of the ten treatments *or* to the baseline. All ten treatments and the baseline were used in Experiment 1.

At the end of the experiment, participants were paid for their decisions. Experimental payments were calculated in experimental dollars based on two randomly chosen interactions within the session. The payments in experimental dollars were subsequently exchanged into the Australian dollar at a rate of 0.5 experimental dollars = 1 Australian dollars (AUD), a rate announced at the start of the experiment. In addition, participants received a show-up fee of 10 AUD, which were earned in addition to the outcome of the two periods randomly selected for payment. On average, participants in the student sessions earned 40.55 AUD [SD = 11.08] throughout the experiment, which lasted approximately 55 min.

Experiment 2

To further understand the applicability of the findings of Experiment 1 in a setting that includes tradespeople, we recruited 87 volunteer non-student participants who were active in sectors with significant potential for providing services and receiving payments through cash transactions.[3] In addition, an equal number of student participants were recruited for these sessions to play the role of a buyer. Sessions were conducted in a computer laboratory at Queensland University of Technology in Brisbane, Australia between 18 October and 25 October 2016. Non-student participants were active in typical trades. The most common occupations were carpenters (23 participants), electricians (7), workers in the construction sector (e.g. tilers, 7), air conditioning and refrigeration specialists (5), and plumbers (3). Their average age was 23.2 years old (despite the similar average to students, several participants in this group were older than the student group); 82.7% were male and 78.2% were Australian nationals.

Extending the design used in Experiment 1, all non-student subjects acted only in the role of a service provider in the experiment, as all treatments were targeted at sellers' compliance behaviour. Treatments in Experiment 2 were introduced on a *within-subject* basis, where participants first made decisions in the *baseline* condition and subsequently in one of the treatments. Hence, all participants are exposed to the baseline *as well as* to *one* of the treatments. Treatment differences therefore describe the within-subjects difference in the baseline and treatment, but the different treatments are introduced between subjects. The student participants who attended these sessions took the role of the customer and were seated in a nearby but physically separate laboratory. As before, we only analyse decisions of service providers. Data from student participants (buyers) in Experiment 2 are consequently not used in the analysis.

Because Experiment 2 was designed to test whether the most central and policy-relevant findings from Experiment 1 are applicable to individuals who commonly encounter opportunities for the type of tax (non-) compliance decisions studied, we only used a subset of the treatments. The treatments selected for Experiment 2 were based on the preliminary analysis of treatments in Experiment 1 that identified treatments with the greatest effect, and confirmation by the Australian Tax Office which of the treatments had significant potential for implementation in the actual practice of a

tax authority. Based on these criteria, Experiment 2 tested treatments of a *higher audit rate, assistance*, *moral suasion*, *infrequent reporting*, and *peer effects seller*.

As in Experiment 1, at the end of the experiment, participants in Experiment 2 were compensated based on a rate that was adapted to reflect the higher opportunity costs of non-student participants. Specifically, we used different show-up fees for student and non-student participants in Experiment 2, where non-students received a show-up fee of 120 AUD while students did not receive a show-up fee. In addition, three randomly selected rounds were paid to participants. Experimental dollars were exchanged into AUD at a rate of 0.5 experimental dollars = 1 AUD, a rate announced at the start of the experiment. On average, non-student (student) participants earned 159.97 (54.18) AUD [SD = 15.76 (13.23)] in Experiment 2, which lasted approximately 60 min.

Analysis

We analyse our data using two compliance measures. First, we calculate the fraction of earned income that was declared as the primary outcome measure, which is commonly used in the tax compliance literature. Since income from non-cash transactions is automatically recorded in the system (except in the infrequent treatments), the tax compliance for these transactions will be 1 (full compliance). In the primary analysis, we want to examine the treatment effects on the overall individual tax compliance, accounting for the decisions made by both sellers and buyers (whether a cash discount was offered or accepted). This variable is advantageous as it provides an overall effect of the treatments on compliance, incorporating all important elements of whether a seller was compliant or not. These elements can include the probability of a cash offer being made, which may be a function of the seller's willingness to evade but may be influenced by the seller's perceptions about finding a buyer who will enable evasion. It also includes buyer decisions to accept cash offers. Finally, it incorporates the final decision to evade among those having the opportunity to do so, as well as decisions of all sellers who either did not offer cash offers (e.g. because they always planned to be compliant) or interacted with a buyer who did not accept the cash option. As some of these underlying motivations are not observed in our data but contribute to the conditional nature of cash offer, we report a combined measure for cash offer acceptance and decisions to evade when given the opportunity. Additionally, we provide an overview that highlights which of these factors are the drivers of compliance and non-compliance decisions.

In the secondary analysis, we will focus on transactions where the participants must make a self-declaration (with the opportunity to evade tax) to assess tax compliance. The

---

[3] Of these non-student participants, 67 completed all parts of the study. Some participants had to be sent home due to overbooking and 3 individuals had to leave early in one of the sessions, providing data points in the experimental part but without specifying their characteristics in the questionnaire.

second compliance measure is the amount of tax revenue lost due to non-compliant behaviour in cash transactions. We employ this additional measure as it is of interest to tax authorities and allows intuitive interpretation of the losses in tax revenue due to the hidden economy and self-declaration. As described in the experimental design, the buyer's acceptance of the seller's price offer is elicited separately from the willingness to offer and accept a cash discount (see Figure A12). Hence, we can tell whether the (cash) transaction would have occurred in the absence of a cash discount. This enables us to construct the tax loss variable by calculating the difference in tax collected from (under) declared cash income and the would-be-declared amount in a non-cash transaction.[4]

As mentioned above, we primarily rely on data on seller decisions, both for Experiment 1 and for Experiment 2. Because we have several income declaration decisions per participant (i.e. two in the infrequent treatments and six in all others) and multiple participants in each session, we use multilevel analysis to derive our results. This analysis corresponds to random-effects models for longitudinal data and allows us to control for repeated observations by individual and by session through estimating an individual-level and a session-level random effect. Random-effects regressions allow the greatest control for repeated observations, while (other than fixed-effects regressions) allowing the study of between-subjects (Experiment 1) and within-and-between-subjects (Experiment 2) effects. The numbers of participants (sellers) in each treatment condition for the two experiments are summarised in Table A1. In Table A2, we outline the treatments involved in the first and second part of each experimental session.

As mentioned, student participants switched roles (buyers or sellers) in the second part of the experiment, meaning that the participants only experienced one treatment as a seller. The treatments within a session were designed such that decisions made by the sellers in the second part of the experiment would not be affected by their role in the first part as buyers. For example, the *socially responsible buyer* treatment was only conducted in part 2 (see sessions 10–12).[5] For non-student sessions, participants always play the role of seller throughout the entire experiment, with the first part serving as the baseline.

The data generated by experimental decisions were loaded into the Stata statistical software for analysis (version 16.1). The decision variables of the experiment can be summarised as follows: the average price offered by the seller was 65.2 experimental dollars (standard deviation = 15.4), the frequency of cash offers was 57.9%, and the average declared amount was 48.8 experimental dollars ($SD = 22.1$) for service provider decisions. For customers, 78.2% of all offers were accepted, and the frequency of customer acceptance of cash offers was 83.4%. Of all occurring transactions, 53.1% were cash transactions. These numbers demonstrate a significant scope for evasion, given cash offers were frequently used and substantial amounts were evaded.

It is important to note that caution is necessary when applying absolute levels observed in the laboratory. The absolute values reflect decisions inside the laboratory and similar levels in the real world would be coincidental, i.e. levels in the real world may be substantially different. This includes potential level differences between student and non-student participants. In the experiment, the overall compliance rate of non-student participants is higher (see Table A3 for a summary of decisions by experiment), which is consistent with prior findings that student participants are less compliant in tax evasion games (see meta-analysis results by Alm & Malézieux, 2021; Gërxhani & Schram, 2006; Choo et al., 2016). However, this does not indicate that tax compliance of service providers is higher than the compliance rate of students. What is informative, however, are the differences between the different treatments within the groups of students and non-students. Our analysis focuses on the difference between the treatments and the control group, as the qualitative effects of treatments or interventions have been shown to be similar in the laboratory and the real world (Kessler & Vesterlund, 2015).

Results

Based on the experimental design, the main variable of interest is the share of income declared by service providers. We go beyond looking at individual compliance decisions by also examining the amount of tax revenue lost due to the availability of non-compliance opportunities (i.e.

---

[4] For example, if a seller only declared 50% of the cash income (as their taxable income), for a transaction with price offer of $80 (i.e. $72 cash income earned due to the 10% cash discount), the amount of tax revenue loss due to under-declaration and cash transaction with a tax rate of 40% is equal to ($80×0.4)—($72×0.5×0.4) = $32—$14.4 = $17.6. $32 is the amount of tax revenue that would have been collected from the transaction if no cash discount is offered or accepted.

[5] To ascertain that the experience of playing the role as 'normal' buyer (i.e. buyers who were not from the *socially responsible buyer* treatment) would in no part influence student participant's compliance decision as seller, we examine whether the compliance outcomes would differ between 'part 1 sellers' and 'part 2 sellers' from

Footnote 5 (continued)

the baseline condition. Given that half of the student participants in the baseline condition are randomly allocated the buyer role in Part 1 of the experiment and the other half in Part 2, we should find no statistically significant difference between their compliance decisions as seller, under the assumption that buyer experience has no effect on compliance. Indeed, we find this to be the case (see Table A8).

Table 1 Differences between the treatments relative to the baseline

a) Experiment 1 (Student)		Tax compliance ratio			Loss revenue		
Treatment conditions	N	M (SD)	Diff	z-stat	M (SD)	Diff	z-stat
Baseline	42	0.67 (0.34)			8.93 (9.26)		
Enforcement paradigm							
Higher audit rate	25	0.90 (0.23)	0.23	−3.06***	2.91 (5.94)	−6.02	2.99***
Endogenous audit	16	0.88 (0.17)	0.21	−2.19**	3.24 (4.08)	−5.69	2.24**
Service paradigm							
Assistance	25	0.84 (0.25)	0.17	−2.12**	4.75 (6.76)	−4.18	1.85*
Positive feedback	16	0.86 (0.20)	0.18	−1.51	3.78 (5.43)	−5.15	1.53
Infrequent reporting	35	0.78 (0.30)	0.11	−1.89*	5.64 (8.37)	−3.28	1.87*
Infrequent pre-filled reporting	17	0.89 (0.22)	0.22	−2.74***	2.87 (5.80)	−6.06	2.79***
Trust/social paradigm							
Audit remorse	17	0.82 (0.27)	0.15	−1.58	5.32 (7.54)	−3.61	1.27
Moral suasion	27	0.82 (0.32)	0.15	−1.89*	5.53 (9.27)	−3.4	1.81*
Socially responsible buyer	23	0.83 (0.29)	0.16	−1.88*	4.88 (8.27)	−4.05	1.81*
Peer effects seller	23	0.84 (0.26)	0.16	−1.95*	4.14 (6.37)	−4.79	1.97**
b) Experiment 2 (Non-student)							
Baseline	86	0.88 (0.21)			3.53 (5.89)		
Enforcement paradigm							
Higher audit rate	18	0.84 (0.23)	−0.04	0.49	3.77 (5.22)	0.25	−0.36
Service paradigm							
Assistance	18	0.94 (0.15)	0.06	−1.76*	1.77 (4.07)	−1.75	1.77*
Infrequent reporting	15	0.73 (0.33)	−0.15	1.58	8.22 (10.02)	4.69	−1.62
Trust/social paradigm							
Moral suasion	17	0.87 (0.21)	−0.01	0.16	4.17 (6.06)	0.64	−0.4
Peer effects seller	18	0.93 (0.22)	0.05	−2.01**	2.29 (6.33)	−1.24	1.98**

z-statistics from non-parametric Wilcoxon rank-sum test. *, **, and *** indicate 10%, 5%, and 1% level of significance, respectively

self-declaration), which is of crucial interest for the tax administration (how much tax money comes in). We first focus on our results using student participants in Experiment 1 before scrutinising our findings by analysing results for non-student participants in Experiment 2.

Figure 2 shows the mean of a seller's average tax compliance ratio, separated by treatments (with different colours representing *enforcement*, *service*, and *trust/social* paradigms). As can be seen, compliance appears to be higher in all treatments compared to the baseline condition (on average, the compliance rate is 67% in the baseline condition), with some variation between the different treatments. This is particularly true in treatments under the *enforcement paradigm*, as the tax compliance rate increased on average by 21 (*endogenous audits*) and 23 (*higher audit rate*) percentage points. On average, the compliance ratio increased by 15–16 percentage points in the four treatments under the *trust/social paradigm*. For the *service paradigm* treatments,

we observe a relatively large variation with the treatment effects ranging from 11 to 22 percentage points.

Table 1 (panel a) shows that the differences between the baseline and the relevant treatment groups are statistically significant except for the *positive feedback* and *audit remorse* treatments. We observe in Fig. 3 that the average amount of tax revenue lost to self-reported cash income was significantly lower in most treatment conditions compared to the baseline, where 8.93 experimental dollars were lost per transaction, on average.[6] This demonstrates that most treatments are effective in increasing tax revenue collection as the average tax income loss is reduced to amounts

---

[6] In the absence of cash transactions (i.e. if all transactions occurred were recorded automatically without cash discounted), the average tax revenue per transaction in the baseline condition would amount to 25 experimental dollars. This indicates that 36% of total tax revenue was lost due to cash income reporting.

Table 2 Multilevel random-effects regressions on tax compliance (Student)

	A	B	C	D	E	F
	Dep. Var. = Compliance ratio			*Dep. Var. = Tax income loss*		
Enforcement paradigm						
Higher audit rate	0.23***	0.24***	0.24***	−6.06***	−6.35***	−6.08***
	(4.63)	(4.63)	(4.25)	(−4.17)	(−4.50)	(−3.93)
Endogenous audit	0.21***	0.19***	0.18***	−5.71***	−5.27***	−4.99***
	(6.24)	(8.17)	(7.28)	(−5.94)	(−7.26)	(−7.22)
Service paradigm						
Assistance	0.18**	0.17**	0.19**	−4.25*	−4.19**	−4.75**
	(1.99)	(2.23)	(2.43)	(−1.73)	(−1.96)	(−2.35)
Positive feedback	0.19***	0.22***	0.19***	−5.24***	−6.05***	−5.02***
	(3.13)	(3.36)	(2.88)	(−3.28)	(−3.73)	(−2.94)
Infrequent reporting	0.11**	0.11***	0.11**	−3.24**	−3.35**	−3.08**
	(2.16)	(2.65)	(2.44)	(−2.07)	(−2.54)	(−2.25)
Infrequent pre-filled reporting	0.23***	0.20**	0.20**	−6.16***	−5.47**	−5.29***
	(3.26)	(2.52)	(2.57)	(−3.17)	(−2.52)	(−2.64)
Trust/social paradigm						
Audit remorse	0.15***	0.17***	0.18***	−3.66***	−4.09***	−4.47***
	(4.66)	(3.50)	(3.77)	(−3.36)	(−2.96)	(−3.44)
Moral suasion	0.15***	0.14***	0.16***	−3.51***	−3.26***	−3.89***
	(3.92)	(5.96)	(6.39)	(−2.61)	(−3.82)	(−4.74)
Socially responsible buyer	0.16***	0.18***	0.19***	−4.10***	−4.66***	−4.94***
	(5.45)	(11.53)	(11.23)	(−3.57)	(−6.50)	(−7.85)
Peer effects seller	0.17***	0.16***	0.15***	−4.81***	−4.66***	−4.08***
	(4.54)	(7.21)	(6.53)	(−4.80)	(−7.04)	(−7.67)
Controls						
Male		−0.10***	−0.10***		2.71***	2.74***
		(−3.53)	(−3.46)		(3.89)	(3.75)
Age		0.01***	0.01***		−0.29***	−0.29***
		(3.43)	(3.22)		(−3.77)	(−3.37)
Domestic		0.10***	0.10***		−2.99***	−2.77***
		(3.83)	(3.60)		(−4.13)	(−3.91)
Price offered			−0.002***			0.11***
			(−3.82)			(7.60)
Constant	0.67***	0.53***	0.68***	8.99***	12.94***	5.58**
	(26.12)	(7.10)	(6.31)	(9.71)	(6.84)	(2.14)
N	1241	1241	1241	1241	1241	1241
Log pseudolikelihood	−13.814	−2.703	2.770	−4088.9	−4077.5	−4059.8

Baseline condition serves as the reference group. Standard errors are clustered at the session level. *, **, and *** indicate 10%, 5%, and 1% level of significance, respectively

between 2.87 and 5.64 experimental dollars.[7] Nevertheless, the differences between *positive feedback* and *audit remorse* treatments compared to the baseline were not statistically significant (Table 1 panel a). In addition, while on average

the treatments under the *enforcement* paradigm have higher levels of compliance (higher compliance ratio and tax revenue lost to self-reported cash income), we do not find a statistically significant difference to the outcome observed in treatments under the *service* or *trust/social* paradigms, nor are the differences statistically significant between the latter two paradigms (Figs. 2 and 3).

In Fig. 4, we report the average declared-earned income ratio (left panel) and loss of tax revenue (right panel) by

---

[7] For comparison, tax revenue collected for each fully compliant transaction amounts to about 26.8 experimental dollars. This translates to a loss of 10.7 to 21.1% in tax income.

Table 3 Predicted mean treatment differences—Student session

Tax compliance ratio	Mean	Std. err	Lower bound	Upper bound
Enforcement paradigm	0.224	0.006	0.211	0.237
Higher audit rate	0.23	0.008	0.214	0.245
Endogenous audit	0.215	0.011	0.192	0.238
Service paradigm	0.168	0.005	0.159	0.178
Assistance	0.181	0.007	0.168	0.194
Positive feedback	0.192	0.007	0.178	0.207
Infrequent reporting	0.115	0.006	0.103	0.127
Infrequent pre-filled reporting	0.237	0.015	0.208	0.266
Trust/social paradigm	0.163	0.004	0.156	0.171
Audit remorse	0.156	0.01	0.136	0.175
Moral suasion	0.157	0.006	0.144	0.169
Socially responsible buyer	0.17	0.007	0.157	0.183
Peer effects seller	0.171	0.009	0.153	0.19
Tax revenue loss				
Enforcement paradigm	−6.015	0.202	−6.414	−5.617
Higher audit rate	−6.044	0.255	−6.549	−5.539
Endogenous audit	−5.972	0.332	−6.633	−5.31
Service paradigm	−4.592	0.147	−4.882	−4.302
Assistance	−4.387	0.191	−4.766	−4.007
Positive feedback	−5.408	0.281	−5.968	−4.848
Infrequent reporting	−3.387	0.205	−3.791	−2.982
Infrequent pre-filled reporting	−6.454	0.446	−7.34	−5.567
Trust/social paradigm	−4.127	0.120	−4.364	−3.891
Audit remorse	−3.678	0.28	−4.235	−3.121
Moral suasion	−3.591	0.19	−3.966	−3.215
Socially responsible buyer	−4.277	0.194	−4.661	−3.892
Peer effects seller	−5.016	0.286	−5.583	−4.448

95% confidence interval of the mean treatment difference

Fig. 2 Tax compliance ratio in student session. Bars represent mean tax compliance ratio of individual averages in each treatment condition. Error bars represent 95% confidence intervals of the mean. Orange, blue, and yellow bars represent treatments under the *enforcement*, *service*, and *trust/social* paradigms, respectively. Mean differences across paradigms are not statistically significant (*n.s.*)
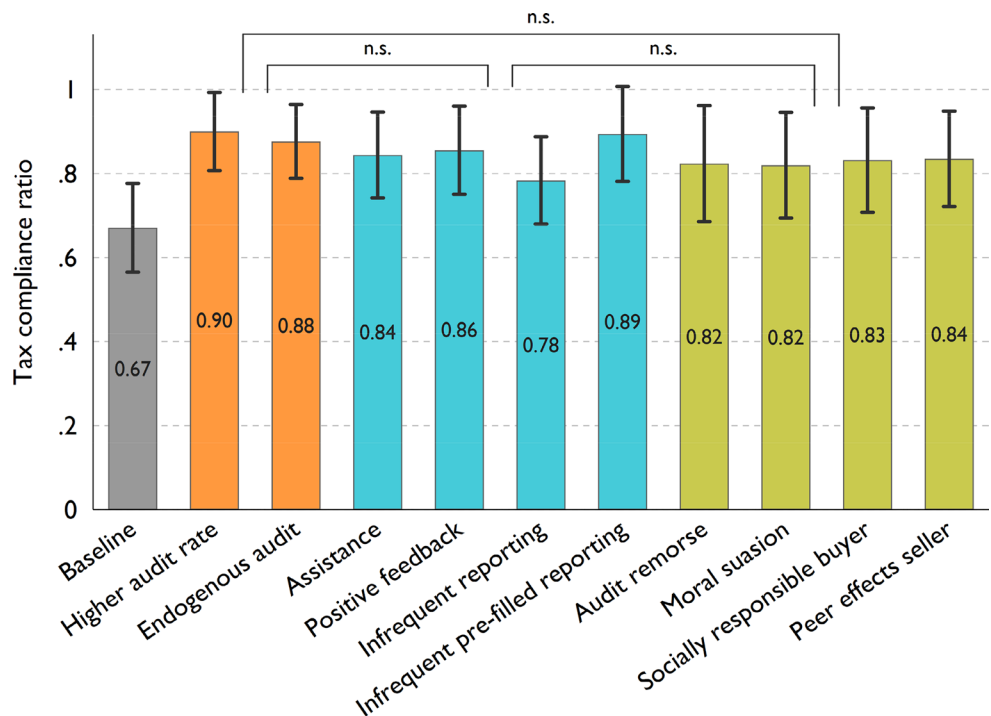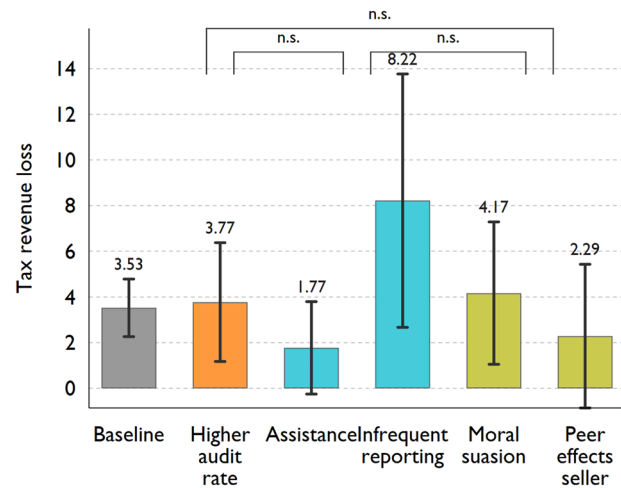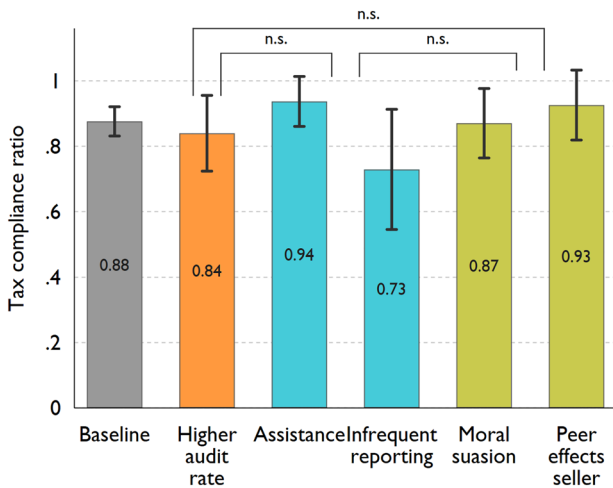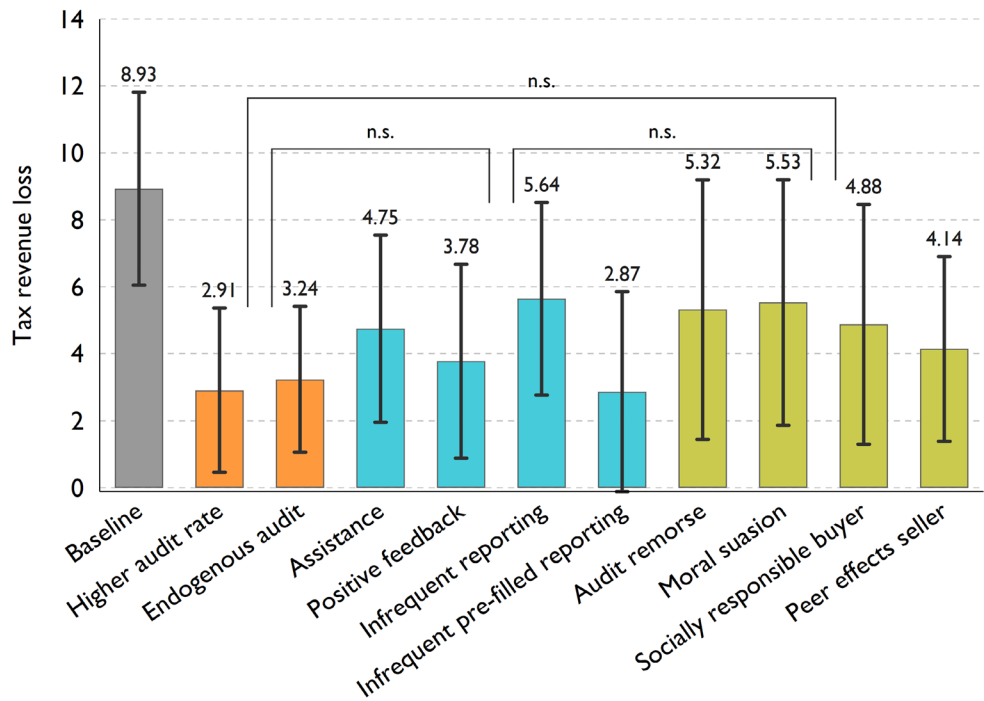
Fig. 3 Loss of tax revenue in student session. Bars represent the mean tax revenue loss of individual averages in each treatment condition. Error bars represent 95% confidence intervals of the mean. Orange, blue, and yellow bars represent treatments under the *enforcement*, *service*, and *trust/social* paradigms, respectively. Mean differences across paradigms are not statistically significant (*n.s.*)

Fig. 4 Tax compliance ratio and loss of tax revenue in non-student session. Bars represent the mean tax revenue loss of individual averages in each treatment condition. Error bars represent 95% confidence intervals of the mean. Orange, blue, and yellow bars represent treat-

ments under the *enforcement*, *service*, and *trust/social* paradigms, respectively. Mean differences between paradigms are not statistically significant (*n.s.*)

treatment for non-student participants (Experiment 2). While the levels of individual tax compliance in the baseline for non-student participants are already high, we find a statistically significant effect in enhancing compliance and reducing tax revenue loss for the *assistant* and *peer effects seller* treatments. In particular, the positive peer effects with other service providers and tax-reporting assistance on compliance indicate that for non-student participants, the ability to compare and receive cooperative help plays a significant role. Nonetheless, Table 1 (panel b) indicates that other treatments are less able to affect tax compliance for non-student participants, as the effects were not statistically significant. Moreover, we observe that *infrequent reporting* leads to lower levels of compliance and results in a larger amount of tax revenue loss for non-student participants, although the effects are not statistically significant. Similar to the student session (Experiment 1), we do not find statistically

significant differences in the mean compliance outcomes between the treatments under the three paradigms (Fig. 4).

Next, we present the multivariate analysis of the treatment effects (Table 2), accounting for the longitudinal structure of the data (repeated individual observations; Model A and D) and effects from other covariates such as demographics, Australian nationality status of the participant (Model B and E), and price offered to the consumer (Model C and F).

For the student session (Experiment 1), we again find that the average treatment effects in all treatment conditions are positive and highly statistically significant, while participant demographics, domestic status, and price offer have a statistically significant effect on individual compliance decisions.[8] However, controlling for these factors does not substantially impact the estimates of the treatment effects. In sum, this indicates that the treatments induce a significant increase in tax compliance in terms of income declaration and reduce tax revenue loss with a size comparable to results reported in Table 1 (see Appendix Figure A14 illustrating effect size differences between each pair of treatment conditions, estimated from the full model specifications C and F in Table 2).[9] Overall, we do not find that the effects of the nine treatment conditions significantly differ from each other, except that the effect from the *infrequent reporting* treatment is slightly weaker than others (at 10% statistical significance compared to the treatments under the *enforcement* paradigm and *socially responsible buyer* treatment). Moreover, while the effects of the treatments under the *enforcement* paradigm are among the largest across all treatments, we do not find any statistically significant differences between the pooled treatment effects across the three paradigms (Table A10 panel b).

In addition, when looking at the predicted mean differences between the treatments and the *baseline* (Fig. 5) it appears that the *enforcement* paradigm delivered the largest effect on increasing tax compliance, while treatments under the *service* and *trust/social* paradigms seem to have smaller effects (Table 3). When testing for the differences of individual predicted compliance (adjusted for effect from all covariates) between the three paradigms, we find that *enforcement-based* approaches further increase the compliance ratio by 5.5 and 6 percentage points relative to *service* and *trust/social* paradigms, respectively (significant at 1%

level with Bonferroni correction). Furthermore, while *infrequent reporting* has the smallest effect of all treatments, *prefilling* the income, even when charging a small economic cost, is highly effective in increasing compliance and reducing tax income losses.

We further extend this analysis by examining whether a *service* or *trust/social* paradigm could be an appropriate substitute for *enforcement*, testing the effect of the *service* and *trust/social* paradigms against enforcement as the baseline (see Table A3). This allows to test our hypothesis regarding the relative effectiveness of the different paradigms. We did not identify statistically significant differences between any of the *service* or *trust/social* treatments and the enforcement treatments, except for the infrequent reporting treatment, which increased revenue loss and reduced compliance behaviour. Indicating that while other interventions can increase compliance, they are not as effective in comparison to *enforcement*.

Next, we investigate the treatment effects using additional controls in the sample of non-student participants (Experiment 2). Table 4 shows the results from the multivariate analysis, using the same estimation approach as employed for student participants via introducing further control variables. By accounting for those factors, treatment effects are more precisely estimated. The estimated treatment differences (relative to the baseline) for non-students are also illustrated graphically in Fig. 6.[10] As can be seen, only three treatments, namely *higher audit rates*, *assistance,* and *peer effects seller*, exhibit a significant effect that increases individual tax compliance and reduces tax income loss. However, these effects are small compared with the student session (Experiment 1), with at most a 3-percentage point increase in declared income ratio and 1.16 experimental dollar reduction (on average) in tax income loss. On the other hand, a statistically significant negative effect on tax compliance for *moral suasion* and *infrequent reporting* is reported. While the effect size of *moral suasion* is small, *infrequent reporting* lowered the relative amount of earned income declared by 15 percentage points (relative to the baseline) and resulted in tax revenue loss of 4.9 experimental dollars on average.

Note that the net effects of these treatments differ to the student session, but the rank order of the treatment effects on compliance is similar in both experiments for students and non-students, as *moral suasion* and *infrequent reporting* also had the smallest effect in the student session. Overall, we find the enforcement approach (i.e. *higher audit rate*) is

---

[8] In particular, the estimated effects of participant's age and gender on compliance aligns with those found in most tax evasion games in prior literature (Alm and Malézieux 2021).

[9] In addition, we also perform the Wald test to assess the joint significance of all treatment effects and effects of treatment under each paradigm (based on Model C and F, see Table A9) as well as conduct regressions where we estimate the treatment effect of the paradigms (see Table A10 panel a). Both additional analyses suggest the overall treatment effects of each paradigm are statistically significant.

---

[10] Pairwise comparisons of all treatment effects in the non-student session are presented in Figure A15. In sum, it shows that the effect sizes of each treatment differ statistically significantly from one another.
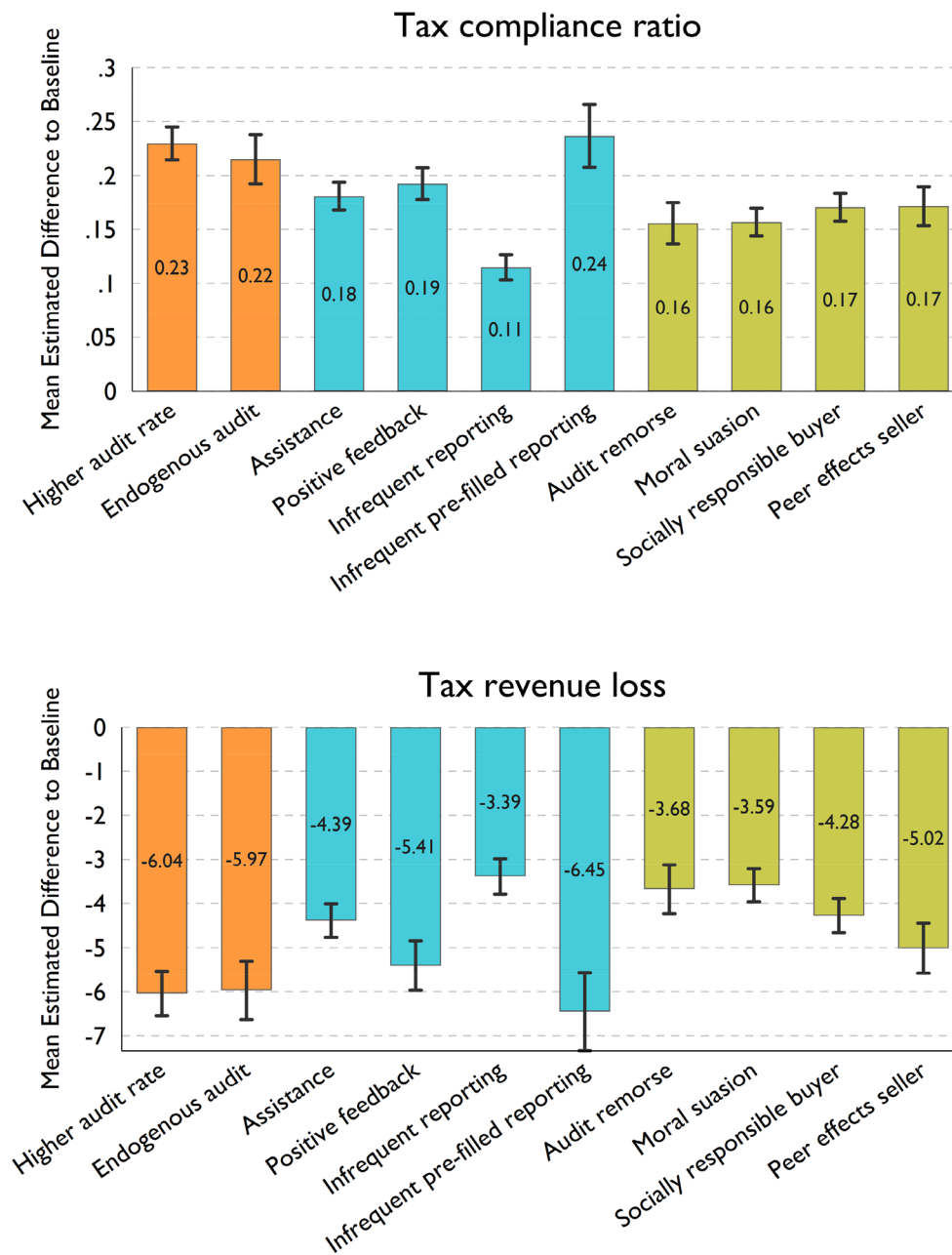
Fig. 5 Predicted treatment-baseline differences of tax compliance ratio and tax revenue loss (Student). Control group is equal to *Baseline*. Error bars represent 95% confidence intervals of the estimated mean difference, computed from Model C and F in Table 2

more effective in increasing tax compliance of non-student participants, compared to treatments under the *trust/social* paradigm (and *service* paradigm, although the difference is not statistically significant, see Table A10 panel b). Nevertheless, when testing the differences of *predicted outcomes* (accounting for the effect from other covariates), we find that enforcement (i.e. *higher audit rate*) results in the highest

tax compliance ratios and least tax revenue losses, while the *service* and *trust/social* paradigm produces an overall net decrease in tax compliance ratios and net increase in tax revenue losses due to the results from *moral suasion* and *infrequent reporting* treatments (see Table 5).

Nonetheless, due to the differences in experimental design in Experiment 1 and Experiment 2 (between-subject

Table 4 Multilevel random-effects regressions on tax compliance (Non-student)

	A	B	C	D	E	F
	Dep. Var. = Compliance ratio			*Dep. Var. = Tax income loss*		
Enforcement paradigm						
Higher audit rate	0.03***	0.03***	0.03***	−0.97***	−0.97***	−1.16***
	(6.21)	(5.39)	(5.05)	(−9.14)	(−8.16)	(−6.71)
Service paradigm						
Assistance	0.03***	0.03***	0.03***	−0.77***	−0.79***	−0.97***
	(8.07)	(7.07)	(6.93)	(−8.76)	(−7.95)	(−6.58)
Infrequent reporting	−0.16***	−0.16***	−0.15***	4.87***	4.86***	4.47***
	(−39.30)	(−37.79)	(−26.37)	(51.78)	(49.98)	(23.52)
Trust/social paradigm						
Moral suasion	−0.02***	−0.03***	−0.02***	0.80***	0.85***	0.34*
	(−6.57)	(−7.98)	(−3.51)	(8.94)	(9.47)	(1.69)
Peer effects seller	0.01***	0.02***	0.02***	−0.20**	−0.22**	−0.44***
	(3.50)	(3.39)	(3.58)	(−2.01)	(−2.08)	(−2.65)
Controls						
Male		−0.03	−0.03		1.10	1.12
		(−1.05)	(−1.11)		(1.02)	(1.15)
Age		0.001	0.002		−0.01	−0.04
		(0.66)	(1.07)		(−0.11)	(−0.76)
Domestic		−0.01	−0.01		−0.34	0.05
		(−0.21)	(−0.40)		(−0.27)	(0.04)
Price offered			−0.001**			0.08***
			(−1.99)			(4.07)
Constant	0.88***	0.92***	0.99***	3.37***	1.77	−2.93
	(39.95)	(12.80)	(9.93)	(6.16)	(0.76)	(−1.09)
N	810	810	810	810	810	810
Log pseudolikelihood	123.887	124.194	126.273	−2555.5	−2555.3	−2544.9

Baseline condition serves as the reference group. Standard errors are clustered at the session level. *, **, and *** indicate 10%, 5%, and 1% level of significance, respectively
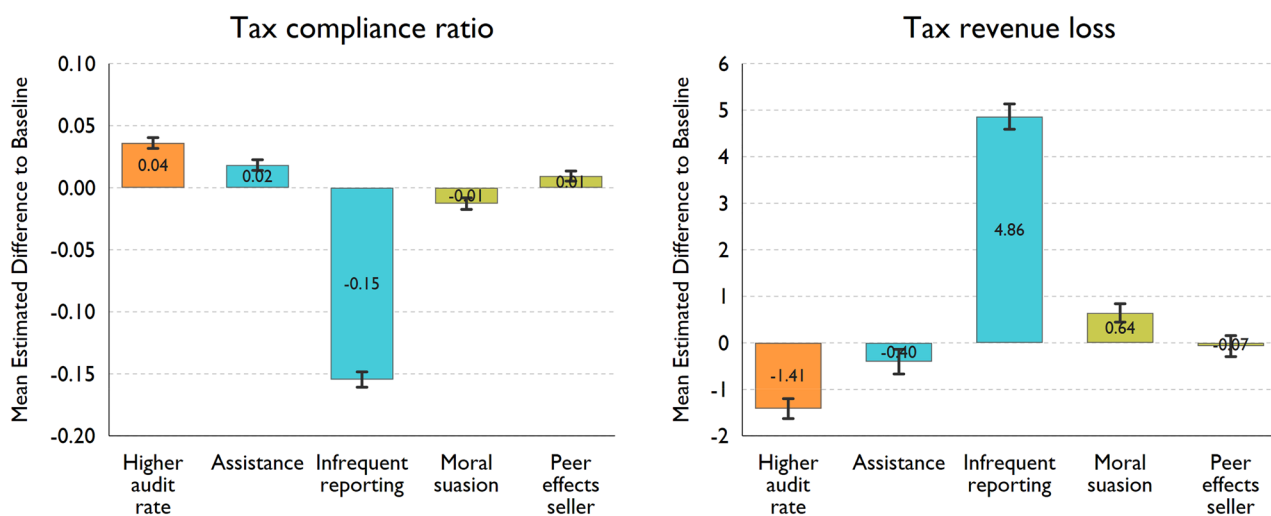

Fig. 6 Predicted treatment-baseline differences of tax compliance ratio and tax revenue loss (Non-student). Control group is equal to *Baseline*. Error bars represent 95% confidence intervals of the estimated mean difference, computed from Model C and F in Table 3

Table 5 Predicted mean treatment differences—Non-student session

Tax compliance ratio	Mean	Std. err	Lower bound	Upper bound
Enforcement paradigm (Higher audit rate)	0.036	0.002	0.032	0.04
Service paradigm	−0.065	0.007	−0.079	−0.051
Assistance	0.018	0.002	0.014	0.023
Infrequent reporting	−0.155	0.003	−0.161	−0.148
Trust/social paradigm	−0.001	0.002	−0.005	0.002
Moral suasion	−0.013	0.002	−0.017	−0.008
Peer effects seller	0.009	0.002	0.005	0.013
Tax revenue loss				
Enforcement paradigm (Higher audit rate)	−1.415	0.108	−1.629	−1.2
Service paradigm	2.134	0.226	1.688	2.580
Assistance	−0.404	0.133	−0.669	−0.138
Infrequent reporting	4.862	0.137	4.59	5.135
Trust/social paradigm	0.280	0.080	0.122	0.438
Moral suasion	0.643	0.099	0.446	0.84
Peer effects seller	−0.07	0.114	−0.296	0.156

95% confidence interval of the mean treatment difference

vs within-and-between-subject design) and sample characteristics between the two sessions, this should primarily be regarded as a qualitative replication of the effects found for students (e.g. the relative ranking of the treatment effects in the two experiments). For example, the additional seller experience (part 1 baseline) for non-student participants might affect their compliance decisions in the treatment conditions (part 2), whereas student participants only played the role of seller once.

We also consider if either cash offers being made by sellers, acceptance of cash offers by buyers, or compliance rates among those who have the opportunity to evade are drivers of the observed overall effect of greater compliance in the treatments relative to the baseline. We do so based on multilevel models as before. We find that both student and non-student sellers are less likely to make cash offers across our different treatment conditions than they do in the baseline. Student sellers reduce the probability of a cash offer by between 8 and 22 percentage points relative to the baseline (the overall average rate of cash offers is 60.44%), and most of these reductions are statistically significant. By comparison, there is no unambiguous effect indicating higher or lower rates of cash offers across treatments relative to the baseline among non-student sellers (see Tables A8). We do not find indications that the change in overall compliance is driven by buyer decisions for either the student or non-student experiment,

as acceptance rates of cash offers are very high and similar across all treatment conditions and the baseline (see Table A9). Furthermore, we observe that the treatments change the rates of individuals who are fully compliant or almost fully compliant (i.e. with a compliance rate of 95% or greater) relative to the baseline when they have to declare their income, contributing to our overall effect on compliance. This increase in the rate of full or almost full compliance rates is particularly visible among student subjects, while it appears less clearly for the non-students (see Table A6).

The above analysis based on the overall compliance rate considers the transactions in which taxable income is automatically registered (i.e. sellers do not have the opportunity to evade tax), and could be due either to the seller not offering a cash discount, the buyer refusing to accept a cash discount, or both (see Table A3). However, focussing on transactions where sellers are required to manually declare income (excluding non-cash transactions) may offer additional insights on the sellers' propensity to evade taxes when the opportunity presents itself. Since sellers might have more opportunity to evade taxes, we predicted that policy treatments under the *service* and *trust/social* paradigms such as *assistance* or *moral suasion* would be less effective, while *enforcement* treatments would induce a more profound effect for those transactions where sellers need to manually declare their income.

Table 6 Tax compliance of transactions with self-declared income

Dep. Var. = Compliance ratio	A Student	B	C	D Non-student	E	F
Enforcement paradigm						
Higher audit rate	0.27***	0.28***	0.28***	0.02***	0.02**	0.02**
	(5.27)	(4.49)	(4.51)	(2.68)	(2.02)	(2.14)
Endogenous audit	0.21***	0.18***	0.18***			
	(7.91)	(8.17)	(7.29)			
Service paradigm						
Assistance	0.14	0.13	0.14	−0.02***	−0.02***	−0.01
	(0.94)	(0.98)	(1.04)	(−4.12)	(−3.74)	(−0.98)
Positive feedback	0.19*	0.22*	0.21*			
	(1.72)	(1.88)	(1.81)			
Infrequent reporting	0.23***	0.23***	0.23***	−0.12***	−0.12***	−0.12***
	(5.55)	(6.78)	(6.62)	(−21.13)	(−19.17)	(−10.59)
Infrequent pre-filled reporting	0.34***	0.30***	0.30***			
	(5.11)	(3.77)	(3.83)			
Trust/social paradigm						
Audit remorse	0.20***	0.21***	0.22***			
	(7.40)	(4.13)	(4.36)			
Moral suasion	0.12***	0.10***	0.11***	−0.05***	−0.06***	−0.05***
	(4.87)	(3.99)	(3.72)	(−10.57)	(−16.10)	(−3.16)
Socially responsible buyer	0.14***	0.17***	0.17***			
	(6.02)	(7.02)	(7.08)			
Peer effects seller	0.14***	0.12***	0.12***	0.04***	0.04***	0.05***
	(5.16)	(5.91)	(5.38)	(7.25)	(7.26)	(3.75)
Controls						
Male		−0.11***	−0.11***		−0.06*	−0.06*
		(−2.70)	(−2.67)		(−1.81)	(−1.70)
Age		0.01***	0.01***		−0.00	−0.00
		(3.27)	(3.12)		(−0.75)	(−0.60)
Domestic		0.15***	0.14***		−0.03	−0.03
		(4.15)	(3.96)		(−1.12)	(−1.27)
Price offered			−0.001			−0.001
			(−1.32)			(−0.54)
Constant	0.55***	0.36***	0.45***	0.79***	0.99***	1.03***
	(36.94)	(3.53)	(3.19)	(18.70)	(8.11)	(6.34)
N	800	800	800	434	434	434
Log pseudolikelihood	8.177	18.727	19.626	37.668	38.109	38.615

Baseline condition serves as the reference group. Standard errors are clustered at the session level. *, **, and *** indicate 10%, 5%, and 1% level of significance, respectively

The results reported in Fig. 7 and Table 6 support these predictions. For Experiment 1, we find that enforcement paradigm treatments result in a higher compliance rate for self-reported income compared to treatments under the *trust/social* paradigm (at 5% level of statistical significance) while they are not statistically different from those under the *service* paradigm. Nevertheless, the effect of the *assistance* treatment became statistically insignificant with respect to self-declared income. Similarly, we find that *infrequent pre-filled reporting* significantly improved

tax compliance on cash transactions. For Experiment 2, we also find that the *assistance* treatment is no longer effective in improving tax compliance for non-student participants. While *enforcement* remains as an effective measure, we observe a larger effect size for the *peer effects seller* treatment.

To test whether the individual treatment effect is different between the self-declared income transactions and the overall compliance rate (see Table 2 and 3), we take the overall sample that was used to estimate the overall
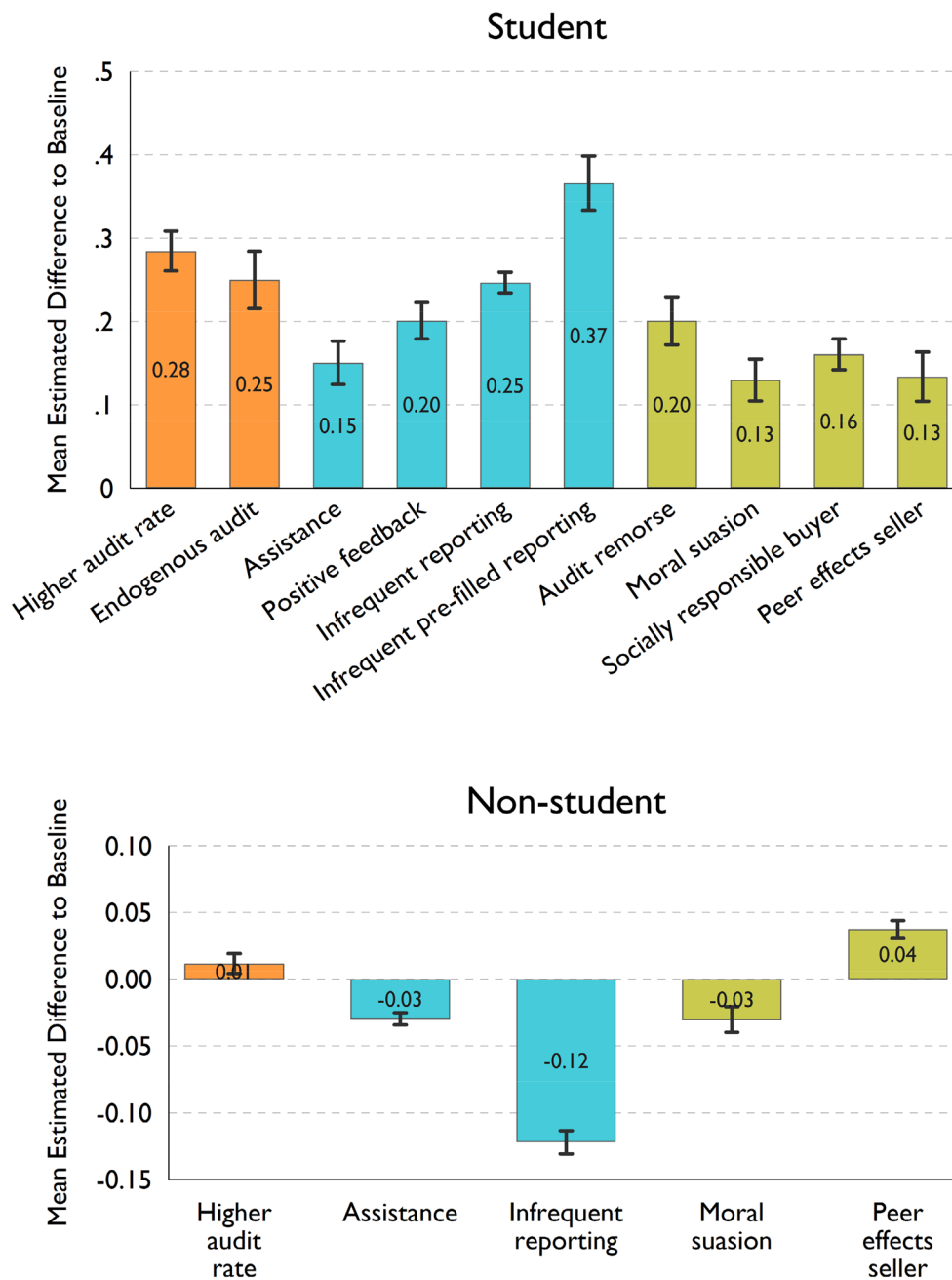
Fig. 7 Predicted treatment-baseline differences of tax compliance ratio of transactions with self-declared income. Control group is equal to *Baseline*. Error bars represent 95% confidence intervals of the estimated mean difference, computed from Model C and F in Table 6

compliance rate by adding the observations of transactions with self-declared income via a dummy variable. Then, we test the interaction terms that involve this dummy variable with the treatment variables (see Table 7). We find that the *higher audit rate* treatment has a stronger positive effect on cash-based income reporting compared to the overall compliance rate for students but has a negative effect for non-students. Interestingly, treatment differences (relative to self-reported income from cash transactions in the

baseline) for *infrequent reporting* increased for cash-based income reporting for both students and non-students. The effects of *assistance* and *moral suasion* seem to have decreased for cash-based income reporting, particularly for non-student participants. *Peer effects seller* is also more effective in improving compliance on cash-based transactions for non-student participants. Overall, when analysing the treatment differences between cash and non-cash-based transactions for each paradigm by pooling treatments, we

Table 7 Tax compliance of transactions with self-declared income

Dep. Var. = Compliance ratio	A Student	B	C	D Non-student	E	F
Interaction terms						
Enforcement paradigm						
Higher audit rate	0.04***	0.04**	0.04**	−0.02***	−0.02***	−0.02***
	(2.65)	(2.13)	(2.50)	(−3.82)	(−4.04)	(−4.16)
Endogenous audit	0.01	−0.00	−0.00			
	(0.52)	(−0.29)	(−0.01)			
Service paradigm						
Assistance	−0.04	−0.04	−0.04	−0.04***	−0.04***	−0.04***
	(−0.62)	(−0.66)	(−0.66)	(−9.51)	(−9.64)	(−5.41)
Positive feedback	0.01	0.00	0.01			
	(0.10)	(0.02)	(0.11)			
Infrequent reporting	0.12***	0.11***	0.11***	0.04***	0.04***	0.04***
	(7.88)	(8.29)	(7.80)	(7.19)	(7.00)	(7.17)
Infrequent pre-filled reporting	0.12***	0.11***	0.11***			
	(7.88)	(8.29)	(7.80)			
Trust/social paradigm						
Audit remorse	0.05***	0.04***	0.04***			
	(3.33)	(3.01)	(2.66)			
Moral suasion	−0.03*	−0.04	−0.04	−0.03***	−0.03***	−0.02***
	(−1.65)	(−1.64)	(−1.43)	(−5.32)	(−5.47)	(−4.80)
Socially responsible buyer	−0.02	−0.01	−0.02			
	(−0.70)	(−0.82)	(−0.80)			
Peer effects seller	−0.02	−0.03*	−0.02	0.03***	0.03***	0.03***
	(−1.10)	(−1.67)	(−1.37)	(4.98)	(5.38)	(4.80)
Controls	No	Yes (demo.)	Yes (all)	No	Yes (demo.)	Yes (all)
N	2041	2041	2041	1244	1244	1244
Log pseudolikelihood	−15.285	6.024	12.560	158.402	159.149	161.974

Listed are estimates of the interaction terms between treatment variables and self-declared income transactions. Standard errors are clustered at the session level. *, **, and *** indicate 10%, 5%, and 1% level of significance, respectively

find that the *service* paradigm is a more effective measure for enhancing cash transaction compliance for student participants, while the *enforcement* paradigm leads to a decrease in compliance for non-student subjects on cash-based income reporting.

Conclusions

The results of the enforcement interventions of *higher audit rate* and *endogenous audit* treatments indicate that deterrence matters for increasing compliance and decreasing tax revenue loss. Despite the greater effect of *enforcement* interventions on compliance and tax revenue, we observe that some interventions founded in the *service* and *trust/social* paradigm can still be a powerful tool for increasing tax compliance in the cash economy context. In fact, a *service* approach can induce significant increases in compliance, as we observe that providing *assistance* maintains a high level of compliance and cooperation of both students and non-students. A similar effect is also observed for non-students using the *peer effects seller* treatment, which demonstrates the importance of social norms. This can be seen as evidence that perceived procedural fairness is a key factor in guaranteeing sustainable compliance.

In comparison to the baseline, the *enforcement*-based approach (*higher audit rate*) has the strongest effect on income declarations in non-students, while cooperation-oriented approaches such as *infrequent pre-filled reporting* or *positive feedback* also improve student declarations. However, for non-students, the *assistance* treatment appears to have a strong effect on tax revenues but less of an effect than the *higher audit rates* treatment. For those who are inclined to use cash transactions, *enforcement* strategies appear to have the strongest influence on tax compliance, as did the other strategies such as *infrequent pre-filled reporting* for students and *peer effects seller* for non-students.

We observed the smallest improvement in tax compliance for student participants in *infrequent reporting* compared to other treatments and even a moderate negative effect in the non-student session, which may be due to the perceived reduction in audit probability or lower psychological costs of evasion, despite giving participants more autonomy. The effect in the non-student session contrasts with the findings of Bérgolo et al. (2017), who found that the reminder of audit probability increased compliance. While we cannot be sure without further research, the effect of infrequent reporting on students compared to non-students may be different due to differing levels of experience in dealing with tax departments. Nevertheless, student compliance greatly improved by introducing the option to pre-populate tax returns (*infrequent pre-filled reporting*) with a relatively small economic cost. This finding aligns with the recent literature on the effectiveness of pre-populated tax returns on compliance (Doxey et al., 2021; Fochmann et al., 2021; van Dijk et al., 2020); that is, pre-filled tax returns can increase compliance, particularly if they are pre-populated correctly and accurately.

Lastly, we observe a difference in the *peer effects seller* treatment effect between students and non-students, which may be driven by tradespeople anticipating a greater cost for violating their group's norms relative to student participants. Future research on evading behaviour within the cash economy should consider whether social norms of the service providers' social group play a role in tax evasion and whether this would explain the difference between student and non-student responses.

In summary, our results may permit some recommendations from a policy perspective. The first is that deterrence works in the context of the cash economy; that is, judging by the (relative) effect size and the consistent results across participant groups, *higher audit rates* and *endogenous audit* were two of the approaches that indicated the safest potential for increasing compliance, at least among the student sample. Second, decreasing the frequency of declarations appeared to lower compliance; this, in turn, suggests that establishing *increased frequencies of reporting* may increase declared amounts. Third, although results were significantly less strong for non-student participants, policies appealing to the *service* and *trust/social* paradigm addressing compliance and revenue loss should not be disregarded in the cash economy. Several previous studies have shown that "soft" measures can have positive effects, have the advantage of simplicity in implementation, and may be comparatively cheap. Based on our results from a broad set of policy measures, higher audit rates, endogenous audit, assistance, and the reduction of reporting frequencies are the most promising avenues through which a tax authority may seek to address compliance in business contexts that offer frequent opportunities for cash transactions but cooperative approaches can be added to the mix if they are sufficiently low cost.

Declarations

Conflict of Interest The authors declare they have no known conflicts of interest.

Ethical Approval All procedures performed in the experiments were in accordance with the ethical standards of the hosting institution.

Informed Consent All participants provided informed consent prior to completing the experiment.

References

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics, 1*(3–4), 323–338. https://doi.org/10.1016/0047-2727(72)90010-2

Alm, J. (1999). Tax compliance and administration. In W. B. Hildreth & J. A. Richardson (Eds.), *Handbook on taxation* (pp. 741–768). Marcel Dekker.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. *International Tax and Public Finance, 19*, 54–77. https://doi.org/10.1007/s10797-011-9171-2

Alm, J. (2019). What motivates tax compliance? *Journal of Economic Surveys, 33*(2), 353–388.

Alm, J., & Malézieux, A. (2021). 40 years of tax evasion games: A meta-analysis. *Experimental Economics, 24*, 669–750. https://doi.org/10.1007/s10683-020-09679-3

Alm, J., & Torgler, B. (2011). Do ethics matter? Tax compliance and morality. *Journal of Business Ethics, 101*(4), 635–651. https://doi.org/10.1007/s10551-011-0761-9

Alm, J., Bloomquist, K. M., & McKee, M. (2012). Comparing student and non-student reporting behavior in tax compliance experiments, *The IRS Research Bulletin, Proceedings of the 2011 IRS Research Conference, Recent Research on Tax Administration and Compliance*, 93–98. Retrieved from https://www.irs.gov/pub/irs-soi/11resconcompexp.pdf.

Alm, J., Cronshaw, M. B., & McKee, M. (1993). Tax compliance with endogenous audit selection rules. *Kyklos, 1*, 27–45. https://doi.org/10.1111/j.1467-6435.1993.tb02405.x

Alm, J., Martinez-Vazquez, J., & Torgler, B. (Eds.). (2010). *Developing alternative frameworks for explaining tax compliance*. Routledge.

Alm, J., McClelland, G. H., & Schulze, W. D. (1992). Why do people pay taxes? *Journal of Public Economics, 48*, 21–48. https://doi.org/10.1016/0047-2727(92)90040-M

Alm, J., McKee, M., & Beck, W. (1990). Amazing grace: Tax amnesties and compliance. *National Tax Journal, 43*(1), 23–37. https://doi.org/10.1086/NTJ41788822

Andreoni, J., Erard, B., & Feinstein, J. S. (1998). Tax compliance. *Journal of Economic Literature, 36*, 818–860.

Batrancea, L. M., Nichita, A., Olsen, J., Kogler, C., Kirchler, E., Hoelzl, E., Weiss, A., Torgler, B., Fooken, J., Fuller, J., Schaffner, M., Banuri, S., Hassanein, M., Alarcón-García, G., Aldemir, C., Apostol, O., Weinberg, D., Batrancea, I., Belianin, A., … M. (2019). Trust and power as determinants of tax compliance across 44 nations. *Journal of Economic Psychology, 74*, 102191. https://doi.org/10.1016/j.joep.2019.102191

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76, 169–217.

Bérgolo, M. L., Ceni, R., Cruces, G., Giaccobasso, M., & Perez-Truglia, R. (2017). Tax audits as scarecrows: Evidence from a large-scale field experiment (No. w23631). *National Bureau of Economic Research*. https://doi.org/10.3386/w23631

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives, 12*, 151–170. https://doi.org/10.1257/jep.12.3.151

Blackwell, C. (2010). A meta-analysis of incentive effects in tax compliance experiments. In J. Alm, J. Martinez-Vazquez, & B. Torgler (Eds.), *Developing alternative frameworks explaining tax compliance* (pp. 97–112). Routledge.

Blumenthal, M., Christian, C., & Slemrod, J. (2001). Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax Journal, 54*, 125–138. https://doi.org/10.17310/ntj.2001.1.06

Bott, K. M., Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2020). You've got mail: A randomized field experiment on tax evasion. *Management Science, 66*(7), 2801–2819. https://doi.org/10.1287/mnsc.2019.3390

Boulding, K. E. (1981). *A preface to grant economics: The economy of love and fear*. Praeger.

Braithwaite, V. (2007). Responsive regulation and taxation: Introduction. *Law & Policy, 29*(1), 3–10. https://doi.org/10.1111/j.1467-9930.2007.00242.x

Chan, S. H., & Song, Q. (2021). Implications of tax audit risk, consequences, aggressive behavior and ethics for compliance. *International Journal of Accounting & Information Management, 29*(5), 823–847. https://doi.org/10.1108/IJAIM-09-2021-0183

Choo, C. L., Fonseca, M. A., & Myles, G. D. (2016). Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment. *Journal of Economic Behavior and Organization, 124*, 102–114. https://doi.org/10.1016/j.jebo.2015.09.015

Cialdini, R. B. (2007). *Influence: The psychology of persuasion*. Harper Collins.

Dong, B., Dulleck, U., & Torgler, B. (2012). Conditional corruption. *Journal of Economic Psychology, 33*, 609–627. https://doi.org/10.1016/j.joep.2011.12.001

Doxey, M. M., Lawson, J. G., & Stinson, S. R. (2021). The effects of prefilled tax returns on taxpayer compliance. *The Journal of the American Taxation Association, 43*(2), 63–85. https://doi.org/10.2308/JATA-18-055

Dulleck, U., Fooken, J., Newton, C., Ristl, A., Schaffner, M., & Torgler, B. (2016). Tax compliance and psychic costs: Behavioral experimental evidence using a physiological marker. *Journal of Public Economics, 134*, 9–18. https://doi.org/10.1016/j.jpubeco.2015.12.007

Erard, B., & Feinstein, J. S. (1994). The role of moral sentiments and audit perceptions in tax compliance. *Public Finance, 49*, 70–89.

Farrar, J., Kaplan, S. E., & Thorne, L. (2019). The effect of interactional fairness and detection on taxpayers' compliance intentions. *Journal of Business Ethics, 154*(1), 167–180. https://doi.org/10.1007/s10551-017-3458-x

Feld, L. P., & Frey, B. S. (2002). Trust breeds trust: How tax payers are treated. *Economics of Governance, 3*(2), 87–99.

Feld, L. P., Frey, B. S., & Torgler, B. (2006). Rewarding honest taxpayers. In H. Elffers, P. Verboon, & W. Huisman (Eds.), *Managing and maintaining compliance* (pp. 45–61). Boom Legal Publishers.

Fisher, R. C., Goddeeris, J. H., & Young, J. C. (1989). Participation in amnesties: The individual income tax. *National Tax Journal, 42*, 15–27. https://doi.org/10.1086/NTJ41788770

Fochmann, M., Müller, N., & Overesch, M. (2021). Less cheating? The effects of prefilled forms on compliance behavior. *Journal of Economic Psychology, 83*, 102365. https://doi.org/10.1016/j.joep.2021.102365

Frey, B. S., & Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics, 35*, 136–159. https://doi.org/10.1016/j.jce.2006.10.006

Gërxhani, K., & Schram, A. (2006). Tax evasion and income source: A comparative experimental study. *Journal of Economic Psychology, 27*(3), 402–422. https://doi.org/10.1016/j.joep.2005.08.002

Górecki, M. A., & Letki, N. (2021). Social norms moderate the effect of tax system on tax evasion: Evidence from a large-scale survey experiment. *Journal of Business Ethics, 172*, 727–746. https://doi.org/10.1007/s10551-020-04502-8

Hallsworth, M. (2014). The use of field experiments to increase tax compliance. *Oxford Review of Economic Policy, 30*(4), 658–679.

Horodnic, I. A., Williams, C. C., Windebank, J., Zaiţ, A., & Ciobanu, C. I. (2021). Explaining consumer motives to purchase in the informal economy. *PLoS ONE, 16*(10), e0258686.

Kahan, D. M. (1998). Social meaning and the economic analysis of crime. *Journal of Legal Studies, 27*, 609–622. https://doi.org/10.1086/468037

Kessler, J., & Vesterlund, L. (2015). The external validity of laboratory experiments: The misleading emphasis on quantitative effects. In G. Frechette & A. Schotter (Eds.), *Handbook of experimental economics* (pp. 392–405). Oxford University Press.

Kirchler, E. (2007). *The economic psychology of tax behaviour*. Cambridge University Press.

Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The ''slippery slope'' framework. *Journal of Economic Psychology, 29*, 210–225. https://doi.org/10.1016/j.joep.2007.05.004

Maciejovsky, B., Schwarzenberger, H., & Kirchler, E. (2012). Rationality versus emotions: The case of tax ethics and compliance. *Journal of Business Ethics, 109*(3), 339–350.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research XLV*. https://doi.org/10.1509/jmkr.45.6.633

McKee, M., Siladke, C. A., & Vossler, C. A. (2018). Behavioral dynamics of tax compliance when taxpayer assistance services are available. *International Tax and Public Finance, 25*(3), 722–756. https://doi.org/10.1007/s10797-017-9466-z

Nuttin, J., & Greenwald, A. G. (1968). *Reward and punishment in human learning*. Academic Press.

Ostrom, E. (2005). Policies that crowd out reciprocity and collective action. In H. Gintis, S. Bowles, R. Boyd, & E. Fehr (Eds.), *Moral sentiments and material interests* (pp. 253–275). MIT Press.

Roth, J. A., Scholz, J. T., & Witte, A. D. (1989). *Taxpayer compliance, Vols. 1 and 2*. Philadelphia: University of Pennsylvania Press.

Schneider, F., & Enste, D. H. (2002). *The shadow economy: An international survey*. Cambridge University Press.

Skinner, B. F. (1938). *The behavior of organism*. Academic Press.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature, 57*(4), 904–954. https://doi.org/10.1257/jel.20181437

Smith, K. W. (1992). Reciprocity and fairness: Positive incentives for tax compliance. In J. Slemrod (Ed.), *Why people pay taxes: Tax compliance and enforcement* (pp. 223–258). University of Michigan Press.

Spicer, M. W., & Becker, L. A. (1980). Fiscal inequity and tax evasion: An experimental approach. *National Tax Journal, 33*(2), 171–175. https://doi.org/10.1086/NTJ41862299

Thorndike, E. L. (1911). *Animal intelligence*. Macmillan.

Thorndike, E. L. (1932). *The fundamentals of learning*. Teachers College, Columbia University.

Torgler, B. (2002). Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. *Journal of Economic Surveys, 16*, 657–684. https://doi.org/10.1111/1467-6419.00185

Torgler, B. (2004). Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Economics of Governance, 5*, 235–253. https://doi.org/10.1007/s10101-004-0077-7

Torgler, B. (2007). *Tax compliance and tax morale: A theoretical and empirical analysis*. Edward Elgar.

Torgler, B. (2013). A field experiment in moral suasion and tax compliance focusing on underdeclaration and overdeduction. *FinanzArchiv: Public Finance Analysis, 69*, 393–411. https://doi.org/10.1628/001522113X675647

Torgler, B. (2016). Tax compliance and data: What is available and what is needed. *Australian Economic Review, 49*, 352–364. https://doi.org/10.1111/1467-8462.12158

Torgler, B. (2022). Behavioral taxation: Opportunities and challenges. *FinanzArchiv/Public Finance Analysis, 78*(1), 5. https://doi.org/10.1628/fa-2022-0003

Torgler, B., & Murphy, K. (2004). Tax morale in Australia: What shapes it and has it changed over time? *Journal of Australian Taxation, 7*(2), 298–335.

Torgler, B., & Schneider, F. (2009). The impact of tax morale and institutional quality on the shadow economy. *Journal of Economic Psychology, 30*, 228–245. https://doi.org/10.1016/j.joep.2008.08.004

Torgler, B., Demir, I. C., Macintyre, A., & Schaffner, M. (2008). Causes and consequences of tax morale: An empirical investigation. *Economic Analysis and Policy, 38*, 313–339. https://doi.org/10.1016/S0313-5926(08)50023-3

Torgler, B., Frey, B. S., & Wilson, C. (2009). Environmental and pro-social norms: Evidence on littering. *The B.E. Journal of Economic Analysis & Policy*. https://doi.org/10.2202/1935-1682.1929

van Dijk, W. W., Goslinga, S., Terwel, B. W., & van Dijk, E. (2020). How choice architecture can promote and undermine tax compliance: Testing the effects of prepopulated tax returns and accuracy confirmation. *Journal of Behavioral and Experimental Economics, 87*, 101574. https://doi.org/10.1016/j.socec.2020.101574

Webley, P., Morris, I., & Amstutz, F. (1985). Tax evasion during a small business simulation. In H. Brandstaetter & E. Kirchler (Eds.), *Economic psychology* (pp. 233–242). Trauner: Linz.

Williams, C., & Kosta, B. (2020). Explaining the cash-in-hand consumer culture in the European home repairs and renovations sector. *Journal of Financial Management of Property and Construction, 25*(3), 387–405.